

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 20:47:41 ; Search time 6644.53 Seconds  
(without alignments)  
11088.547 Million cell updates/sec

Title: US-09-434-837-10\_COPY\_1\_1801

Perfect score: 1801

Sequence: 1 atgcgcagcttcacactccct.....ctatgagtcgcagagccga 1801

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_ov:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_da:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_ph:\*

24: em\_pat:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rtd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1765.2	98.0	24595	6	ATACH5	X00493 Agrobacteri
2	1765.2	98.0	24585	1	BD016312	BD016312 Method of
3	1765.2	98.0	24595	6	E00404	E00404 Ti plasmid
4	1765.2	98.0	24595	6	E00546	E00546 DNA fragmen
5	1765.2	98.0	194140	1	AF242881	AF242881 Agrobacte
6	1760.4	97.7	2591	1	TIPIMS2	K0254 Ti plasmid
7	1605.2	89.1	206479	1	AB016260	AB016260 Agrobacte
8	1603.6	89.0	9931	1	AE007927	AE007927 Agrobacte
9	1603.6	89.0	10377	1	AE009419	AE009419 Agrobacte
10	1602	89.0	5968	1	AF126446	AF126446 Agrobacte
11	1602	89.0	29524	1	ATU237588	AJ237588 Agrobacte
12	1597.2	88.7	4660	1	AB025110	AB025110 Agrobacte
13	1467.6	81.5	10200	1	ATTAIAH	X56185 A. tumefacie
14	1467.6	81.5	15463	1	AVU83987	U83987 Agrobacteri
15	1464.4	81.3	14960	1	AVU83986	U83986 Agrobacteri
16	814	45.2	9577	1	AF126447	AF126447 Agrobacte
17	762.8	42.4	3385	1	AF061780	M61151 Agrobacteri
18	644	35.8	5995	1	ATTUAX	M61151 Agrobacteri
19	358.2	19.9	5516	1	IS4AAM	M91609 Plasmid pTi
20	258.4	14.3	3677	1	PSEIAAMH	M11035 P. syringae
21	254.8	14.1	1937	1	ERWIAAM	L33857 Pantoea agg
22	242.6	13.5	4797	1	PSU04358	U04358 Pseudomonas
23	121.6	6.8	2279	6	AR050109	AR050109 Sequence
24	121.6	6.8	2279	6	I24442	I24442 Sequence 8
25	121.6	6.8	21126	1	RIATL	K03313 Integrated
26	121.6	6.8	21126	6	AR050120	AR050120 Sequence
27	121.6	6.8	21126	6	I04770	I04770 Sequence 94
28	121.6	6.8	21126	6	I24453	I24453 Sequence 19
29	66	3.7	217594	1	AP002086	AP002086 Agrobacte
30	65	3.6	167	1	ATTU6787	U16787 Agrobacteri
31	64.4	3.6	1610	1	ARR271050	AJ271050 Agrobacte
32	52.2	2.9	334108	1	AE016857	AE016857 Pseudomon
33	51.4	2.9	1559	8	AKI18627	AKI18627 Arabidops
34	48.6	2.7	2000	6	AX655393	AX655393 Sequence
35	46	2.6	2000	6	AX655393	AX655393 Sequence
36	43.6	2.4	125020	9	AF429315	AF429315 Homo sapi
37	43.6	2.4	300511	1	AE016775	AE016775 Pseudomon
38	43	2.4	35417	3	AF067936	AF067936 Caenorhab
39	42.4	2.4	125020	9	AF429315	AF429315 Homo sapi
40	41.4	2.3	12733	1	ATTUORF	M60490 Agrobacteri
41	41.4	2.3	163280	2	AC141497	AC141497 Rattus no
42	41.4	2.3	226328	2	AC098160	AC098160 Rattus no
43	41.4	2.3	236068	2	AC121721	AC121721 Rattus no
44	41.4	2.3	240507	2	AC103202	AC103202 Rattus no
45	41.4	2.3	251457	2	AC131820	AC131820 Rattus no

## ALIGNMENTS

RESULT 1

LOCUS ATACH5 24595 bp DNA linear BCT 10-FEB-1999

DEFINITION Agrobacterium tumefaciens Ti plasmid pTi15955-T-DNA region.

ACCESSION X00493 J05108 X00282

VERSION X00493.1 GI:39062

KEYWORDS octopine synthetase; plasmid; synthetase; terminal repeat; unidentified reading frame.

SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)

ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

REFERENCE 1 (bases 1 to 24595)

AUTHORS Barker R.F., Idler K.B., Thompson D.V. and Kemp J.D.

TITLE Nucleotide sequence of the T-DNA region from the Agrobacterium

Pred. No. is the number of results predicted by chance to have a

tumefaciens octopine Ti plasmid pTI15955  
JOURNAL Plant Mol. Biol. 2, 335-350 (1983)  
REFERENCE 2 (bases 602 to 14237)  
AUTHORS Giehlens, U., De Beuckeleer, M., Seurinck, J., Deboeck, F., De Greve, H.,  
Lemmerz, M., Van Montagu, M., and Schell, J.  
TITLE The complete nucleotide sequence of the Ti-DNA of the Agrobacterium  
tumefaciens plasmid pTiAch5  
JOURNAL EMBO J. 3 (4), 835-846 (1984)  
MEDLINE 84207942  
PUBMED 6327292  
REFERENCE 3 (bases 1 to 24595)  
AUTHORS Turk, S.C., Nester, E.W. and Hooykaas, P.J.  
TITLE The virA promoter is a host-range determinant in Agrobacterium  
tumefaciens  
JOURNAL Mol. Microbiol. 7 (5), 719-724 (1993)  
MEDLINE 93225814  
PUBMED 8469115  
REFERENCE 4 (bases 1 to 24595)  
AUTHORS Guevara-Garcia, A., Mosqueda-Cano, G., Arguello-Astorga, G.,  
Simpson, J. and Herrera-Estrella, L.  
TITLE Tissue-specific and wound-inducible pattern of expression of the  
manopine synthase promoter is determined by the interaction  
between positive and negative cis-regulatory elements  
JOURNAL Plant J. 4 (3), 495-505 (1993)  
MEDLINE 94035196  
PUBMED 8220492  
FEATURES  
source  
1. .24595  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:358"  
/plasmid="Ti"  
909..23782  
/note="T-region"  
909..932  
/note="direct repeat A"  
1337  
/note="C is T in [2]"  
/citation=[2]  
1661..2179  
/note="open reading frame 1"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA25163.1"  
/db\_xref="GI:39063"  
/db\_xref="SWISS-PROT:P04028"  
/translation="MYDGPINFIIDSSNLQDRRELKLVLIHTENAYRSSAQRSLIAS  
QRSMINFIINTDVIDPDAKDEYKSCRVACCPDPIPDIDLVSLIYVCSFOEMR  
RYAQRFDGVSOGAVISTVPPYAEGITQTWRMLQKVMQTSKETHDLDAVIAL  
PNIRFKIQISHI"  
complement(1736..2113)  
/note="open reading frame 2"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA25164.1"  
/db\_xref="GI:39064"  
/db\_xref="SPTREMBL:O44387"  
/translation="MCFPACILPPIFUPQPHSLFCYSFGIGMDGDCNCAVGNAYIES  
LICIPIHFKRAINQORYIEYIKGYICWMAARNFARALDHVFVFSWGNINISINDEI  
YPRALRSYETSLRTAISILCVN"  
1948  
/note="C is T in [2]"  
/citation=[2]  
2146  
/note="T is CG in [2]"  
/citation=[2]  
complement(2907..3287)  
/note="open reading frame 3"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA25165.1"  
/db\_xref="GI:39065"  
/db\_xref="SWISS-PROT:P03867"  
conflict  
CDS  
/translation="MNFADTPLASLDLDWACEEFIKTYGASPOLTEGEVLIQTNNGLY  
LVKGSLSQRHIDTHLKFKEKEELSTTIKPKAMKAOQSLTFYVAIPQSNYFLCVSN  
PEKGFURCHNRPRPIYIVAHGMS"  
3369..3370  
/note="TC is TCG in [2]"  
/citation=[2]  
3810..3811  
/note="AA is A in [2]"  
/citation=[2]  
complement(4060..5463)  
/note="open reading frame 4 gene tms"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA25166.1"  
/db\_xref="GI:39066"  
/db\_xref="SWISS-PROT:P03868"  
/translation="MVAITSLAQSLEHLRKDYSCLELVETLIARCEAKSNALLAT  
DMDLRSASAKIDRHGNAGVLCIPLCFKANIATGVPFPTSATPALINHLPIPSRV  
AERLPSAGALPGASGMHELSFGITSNNTATGAVRPNMPPDLPGSSGCVAAVASR  
LMIGIGTDPGASVRLPALCGVYGRPTLGRYPGRITIPVSPTRTPGIACQADV  
VILDRISGTPERIPVPPLKGRIGIPPTYFYDDLDADVAALAEITRIILANGCTPV  
EANIPLHDLNKASFPVALYEPHALKQYLDYFVTSFSDVFKIIRSDVANIANA  
OIDGHSIAEVELARSPRPLQATYRVYFKLNRLDALFPAIPLVARIQDSSVIT  
HNGTMIDTFKIYVRNVDPSSNGLPGLSIPVCITPDLPGMEIDGLAUSDORLAI  
GALEBAIGFRYFAGLPN"  
5561  
/note="T is C in [2]"  
/citation=[2]  
5809..8076  
/note="open reading frame 5 gene tms"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA25167.1"  
/db\_xref="GI:39067"  
/db\_xref="SPTREMBL:O44388"  
/translation="MSASPDLNQCDELPTKAVDLTMVDADELDRRVSDAFLEREAS  
RGRRIQISTECSAGLACKRLADGRPEISAGKVAIVAYIYIGKEILGRILIESKP  
ARATVGLVAIDAPRCMDPSEKOLIQALFLSGKRCAPIDSHPAVISISKAGPRT  
LMPVYENCTMKCTGPTTITLBSGAVPFDVAVGRNMLKSGASFPITDLYVRPFF  
DQSDSGRIGFPEDVPKRVAVIGAGISGLVYANLHAGVDVTTIYASDVGKTL  
WSHAFDAPSVVAEWGAMRFPPAFCFLFFLELYGSSKPPFNPGTVDYIVLYQGV  
YMKKAQQLPPKLFHRVYNGMRAFLKDGFFHERDVLASPAVITOLKSGDIRMAHDSQ  
IMLNRRGSEFSGSGIERIFLGTHPGEFTWSFPHDWDLFLKMGIGSGFGPVEGFI  
EILRIYINGEENQRMCPGISSELPRIASEVYGVVSORICHOVVRALIQEKTKIK  
IRKSGISSELYDQVYVYTSGLANQLBHCITCQNIINIOAPVNOAVDMSHMTGSKPLM  
TERKFWLHLIPSCVIMDGIACAVYCLDBSQDPNKGVLVISTYEDSHKLLAIPD  
KERRLCFLDARSFPAPAGHLFPAQADYDQVNIQHWLTIDENAGAFKXLRNGDQF  
YSEELFEQALDTPANDPNDITISYHKMGFPFGYKPGNAFRAILVQRHMGSGACG  
PLEHSKRYNRN"  
complement(6259..6579)  
/note="open reading frame 6"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA25168.1"  
/db\_xref="GI:39068"  
/db\_xref="SPTREMBL:O44389"  
/translation="MKQPVCHSSGNASANDRFRNLIRKADPSTIGTIYKRSV  
VEQVUCMGTCTTLOHYSTISYHKMGFPFGYKPGNAFRAILVQRHMGSGACG  
SLRD"  
6894..6895  
/note="TC is CT in [2]"  
/citation=[2]  
7597  
/note="T is C in [2]"  
/citation=[2]  
7778..8113  
/note="open reading frame 7"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA25169.1"  
/db\_xref="GI:39069"  
/db\_xref="SPTREMBL:O44390"  
/translation="MLFNMIGLQTRWPGELSNSTGVVRIIFILNFSFKWTRLMIPDF

```
THRAVPSQVDMRVLFBRPVTPSVQLSTIVEAFMQRAILSNLTITALEISLM
ILQVQLV"
conflict
7961
/note="T is A in [2]"
/citation=(2)
7964
/note="A is G in [2]"
/citation=(2)
7968..7969
/note="CG is GC in [2]"
/citation=(2)
8387
/note="A is G in [2]"
/citation=(2)
8771..9493
/note="open reading frame 8 gene tmr"
CDS
/codon_start=1
/translation="MDHLIFGPTCTGKTTTALAQQTGLPVLSDRVQCCQLSTG
SGRPVEELKGTIRLYLIDRPLVEGLIAAQAHRLIEVYHNEANGLLLEGSTSL
LNCMAKNSYSDPFRWHIIRHKLDPDETFMKAARVKQMLHPAGHSIIQELVYDMN
EPLRLPLKEIDGVRVAMLPASONQITADWLQLDANMEGLINGIAQEVFIHAPQOE
OKRPQVMAAFDGEFEGHFGMY"
9564
/db_xref="SWISS-PROT:P03869"
/translation="MDHLIFGPTCTGKTTTALAQQTGLPVLSDRVQCCQLSTG
SGRPVEELKGTIRLYLIDRPLVEGLIAAQAHRLIEVYHNEANGLLLEGSTSL
LNCMAKNSYSDPFRWHIIRHKLDPDETFMKAARVKQMLHPAGHSIIQELVYDMN
EPLRLPLKEIDGVRVAMLPASONQITADWLQLDANMEGLINGIAQEVFIHAPQOE
OKRPQVMAAFDGEFEGHFGMY"
9564
/note="T is C in [2]"
/citation=(2)
9926..9927
/note="TT is T in [2]"
/citation=(2)
9996..10571
/note="open reading frame 9"
/codon_start=1
/translation="MDHSMARQGWLPCLSHGKDQLQGLSELSESKVREKQTDLH
/db_xref="SWISS-PROT:P04030"
/translation="MDHSMARQGWLPCLSHGKDQLQGLSELSESKVREKQTDLH
TKSGDILNPGSEFLYILDKENYRLCRQWVLVSNM5DGLATLLEPSDGYFRQVR
Query Match 98.0%; Score 1765.2; DB 1; Length 24595;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
QY 1 ATGTCAAGCTTCACCTCTCTTGATTAACCAAGTGCATCATCTCCCAACCAAAATGATGAT 60
Db 5809 ATGTCAAGCTTCACCTCTCTTGATTAACCAAGTGCATCATCTCCCAACCAAAATGATGAT 5868
QY 61 CTGACAAATGTCGATTAAGCGGATGAATTGACCGCAGGGTTTCGATGCTTTCTTAGAA 120
Db 5869 CTGACAAATGTCGATTAAGCGGATGAATTGACCGCAGGGTTTCGATGCTTTCTTAGAA 5928
QY 121 CGAAGAGCTTCTAGGGGAAGAGGATTAACCAAGTCCCAACCAAGTGCAGCGCTGGTTA 180
Db 5929 CGAAGAGCTTCTAGGGGAAGAGGATTAACCAAGTCCCAACCAAGTGCAGCGCTGGTTA 5988
QY 181 GCTTGCAAAAGGCTGCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAAAGTAGCA 240
Db 5989 GCTTGCAAAAGGCTGCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAAAGTAGCA 6048
QY 241 GTTCTCTCCGCTTATATCTATATTTGGCAAGAAATTTCTGGGGCGGATATCTTGAATCGAAA 300
Db 6049 GTTCTCTCCGCTTATATCTATATTTGGCAAGAAATTTCTGGGGCGGATATCTTGAATCGAAA 6108
QY 301 CTTTGGGCGGGGCAAGTAGTGTCGTGTTGCCATTCGACTTGGACACCATTTGGATG 360
Db 6109 CTTTGGGCGGGGCAAGTAGTGTCGTGTTGCCATTCGACTTGGACACCATTTGGATG 6168
QY 361 GATTTCCTCCGAAGCACAATTAATCAAGCCCTGTTTTTGTGAGCGGTAAAGATGTGCA 420
Db 6169 GATTTCCTCCGAAGCACAATTAATCAAGCCCTGTTTTTGTGAGCGGTAAAGATGTGCA 6228
```

```
QY 421 CGATTGATCTTAGTCAATTCGNGGCCATTTCAATCTTAAGACTGCGCGCTTTCGAAC 480
Db 6229 CGATTGATCTTAGTCAATTCGNGGCCATTTCAATCTTAAGACTGCGCGCTTTCGAAC 6288
QY 481 CTGCCAATGCGCGTGTACAGAAATGACAGATGAATGCGTTACCGGGTTTACCATTAAC 540
Db 6289 CTGCCAATGCGCGTGTACAGAAATGACAGATGAATGCGTTACCGGGTTTACCATTAAC 6348
QY 541 CTTGAAGGGGCGCTGCCATTTTGAATGATGCTTAATGTCGAAAACCTGATGCTGAAGGGT 600
Db 6349 CTTGAAGGGGCGCTGCCATTTTGAATGATGCTTAATGTCGAAAACCTGATGCTGAAGGGT 6408
QY 601 TCGGCGAGTTCTCTTCCAAACAATCGACTTGTCTAAGACGCGAGCGTTTTTTCACAA 660
Db 6409 TCGGCGAGTTCTCTTCCAAACAATCGACTTGTCTAAGACGCGAGCGTTTTTTCACAA 6468
QY 661 TGTTCGATAGTGAACGATTCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720
Db 6469 TGTTCGATAGTGAACGATTCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 6528
QY 721 GTCATTTGGCGCTGGCATTTTCCGACTCTGTGTGGCAACGAACTGCTTCACTGCTGGGCTA 780
Db 6529 GTCATTTGGCGCTGGCATTTTCCGACTCTGTGTGGCAACGAACTGCTTCACTGCTGGGCTA 6588
QY 781 GACGATGTTAACAATATAGAACGAATGATTCGTTGGAGGCAAGCTTTGGTCAATGCT 840
Db 6589 GACGATGTTAACAATATAGAACGAATGATTCGTTGGAGGCAAGCTTTGGTCAATGCT 6648
QY 841 TTCAGGAGACGCTCTAGTGTCTGTGGCCGAATGSGGGCGATGCGATTTCTCTGCTGCA 900
Db 6649 TTCAGGAGACGCTCTAGTGTCTGTGGCCGAATGSGGGCGATGCGATTTCTCTGCTGCA 6708
QY 901 TTCTGCTTGTTTTCTTCTCTCGACGCTTACGCGCTGTCTTTCGATGAGCGCTTCCCAAT 960
Db 6709 TTCTGCTTGTTTTCTTCTCTCGACGCTTACGCGCTGTCTTTCGATGAGCGCTTCCCAAT 6768
QY 961 CCGGCGACAGTGCACACTTACTTGTGCTCAACAGGGCTCCAAATACATGTGGAAAGCGGG 1020
Db 6769 CCGGCGACAGTGCACACTTACTTGTGCTCAACAGGGCTCCAAATACATGTGGAAAGCGGG 6828
QY 1021 CAGCTCCACCGAAGCGTGTCCATCGCGTTTACAAAGTGGCGCTGCTTCTTGAAGAC 1080
Db 6829 CAGCTCCACCGAAGCGTGTTCATCGCGTTTACAAAGTGGCGCTGCTTCTTGAAGAC 6888
QY 1081 GGTTTTCATGAGCGAGATATGTGTGCTGCTGCTGTGCTTATTACTGACGCTTGA 1140
Db 6889 GGTTTTCATGAGCGAGATATGTGTGCTGCTGCTGTGCTTATTACTGACGCTTGA 6948
QY 1141 TCAAGACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTCCGGGAGG 1200
Db 6949 TCAAGACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTCCGGGAGG 7008
QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGGGACACATCTCTCGTGTGT -AA 1259
Db 7009 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGGGACACATCTCTCGTGTGTGTGAA 7068
QY 1260 ACATGAGATTTTCTCATGATTTGGGACATTAATCAAGTAAAGGATTTGGGCTGCGC -GG 7127
Db 7069 ACATGAGATTTTCTCATGATTTGGGACATTAATCAAGTAAAGGATTTGGGCTGCGC -GG 7127
QY 1320 GTTTGTCAGTTTGAAGAGCGGTTTATGAGATCTCCCTGGTGTGATCAACGAGTA 1379
Db 7128 GTTTGTCAGTTTGAAGAGCGGTTTATGAGATCTCCCTGGTGTGATCAACGAGTA 7187
QY 1380 TGAAGAAATACAGCGGATGTCCTGAAGAAATCTCAGAACTTCCAGTCCGATCCCATC 1439
Db 7188 TGAAGAAATACAGCGGATGTCCTGAAGAAATCTCAGAACTTCCAGTCCGATCCCATC 7247
QY 1440 TGAAGTGTAAACGATGTGTCTGTGAAGCAGCGCATATGCCATGTTCAAGTCAGGCGAT 1499
Db 7248 TGAAGTGTAAACGATGTGTCTGTGAAGCAGCGCATATGCCATGTTCAAGTCAGGCGAT 7307
```

Qy	1500	CCAGAAGAAAGACAAAATTAAGATTAGGCTTAAGGCGGGATATCTGAACCTTTATGA	1559
Db	7308	TCAGAAGAAAGACAAAATTAAGATTAGGCTTAAGGCGGGATATCTGAACCTTTATGA	7367
Qy	1560	TAAAGTGTGTGCATCTGGATCTGCMAATATCCAATCAGGATTTGCTGCATGCGA	1619
Db	7368	TAAAGTGTGTGCATCTGGATCTGCMAATATCCAATCAGGATTTGCTGCATGCGA	7427
Qy	1620	TACCAATATTTTTCAGGCACCAAGTGAACCAAGCGGTGATTAACGCCATATAGACAGATC	1679
Db	7428	TACCAATATTTTTCAGGCACCAAGTGAACCAAGCGGTGATTAACGCCATATAGACAGATC	7487
Qy	1680	GTCAAAACTCTTCTCGATGACTGAACGAAATTTCTGTTAGACCAATCTCCCGTCTTG	1739
Db	7488	GTCAAAACTCTTCTCGATGACTGAACGAAATTTCTGTTAGACCAATCTCCCGTCTTG	7547
Qy	1740	TGTTCTCATGGAACGGATCCGAAAGCAGTGTATTGCTCGACATATGATGTGGCAGATCC	1799
Db	7548	TGTTCTCATGGAACGGATCCGAAAGCAGTGTATTGCTCGACATATGATGTGGCAGATCC	7607
Qy	1800	GA 1801	
Db	7608	GA 7609	

RESULT 2			
BD016312	24595 bp	DNA	linear
LOCUS	BD016312		PAT 27-AUG-2002
DEFINITION	Method of promoting plant transcription by using octopine T-DNA		

ACCESSION	BD016312
VERSION	BD016312.1
KEYWORDS	GI:22557450
SOURCE	UP 2001190289-A/1
ORGANISM	Agrobacterium tumefaciens (Rhizobium radiobacter)

REFERENCE	1	2	3
AUTHORS	1 (bases 1 to 24595)		
TITLE	Barker, R. F. and Kemp, J. D.		
JOURNAL	Method of promoting plant transcription by using octopine T-DNA		
	Patent: JP 2001190289-A 1 17-JUL-2001;		

COMMENT	OS	Agrobacterium tumefaciens
	PN	JP 2001190289-A/1

PF	22-NOV-2000	JP	2000356816	
PR	18-NOV-1983	US	553786	
PI	RICHARD F BARKER, JOHN D KEMP			
PC	C12N15/09, A01H5/00, C12N5/00, C12N15/00, C12N5/00	CC		Method of
	promoting plant transcription by using octopline T- CC			
	DNA promoter			
PH	Key			
	Location/Qualifiers			
PT	source		1. 24595	
PT	/Organism:'Agrobacterium tumefaciens'.			

FEATURES	Location/Qualifiers
SOURCE	1. .24595

BASE COUNT	6534	a	5509	c	5793	g	6759	t
ORIGIN								

Query Match	98.0%	Score 1765.2;	DB 6;	Length 24599;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1792;	Conservative	0;	Mismatches 8;	Indels 2;
				Gaps 2;

1 ATGTCACTTCACCTCTCCCTGTAACCAAGTGGATCATCTCCACCAAAATGTGGAT 60

61 CTGACAAATGGTTCGATAAAGGCCGATGCAATTGGACCGCAGGGTTTCGATCGCTTCTTAGAA 120

Db 5869 CTGACATGGTCGATTAGGCGGATGATTGACCGCAGGGTTCCGATGCTTCTTAGAA 5928

121 CGAGAGCTTCTAGGGGAGGAGTACTCAAGCTCCACCGAGTCAGCGCTGGGTTA 180

DB 3929 CGAGAGCTTCTAGGGGAGAGAGATATCTATATCTCCATCCGAGATGCAAGGCTGGGATA 3988

Db 5989 GCTTGCAAAGCGTGGCCGATGTCGCTTCCCGAGATCTCAGCTGCTGGAAAGTACGA 6048

241 GTTCTCTCCGCTATATCTATATTGGCAAGAAATTCTGGGGCGGATCTTGAATCATA 300

[illegible]

Db 6109 CCTGGGGCGGGGCAACAGTGTGCTCTCGTTGCCATCGACTTGGCACCATTTCGATG 6168

361 GATTTCCTGGAGCACACATACTCAAGCCCGTATTCGTCGAGCGGTATTTGATGTGCA 420

421 CCGATTGATCTTAGTCATTTCTGCGCCATTTCAATCTCTAAGACTGCCGGCTTTCGAAC 480

db 6229 CCGATTGATCTTAGTCAITTCGTGCGCAITTCATCTCTAAGACTGCCGGCTTCCGAAC 6288

[illegible]

541 CTTGAAGGGGCCGTGCCATTGACATGCTTATGGTCGAAACCTGATGCTGAAGGT 600

6349 CTTGAGGGGCGGTCATTTGACATGGTACCTATGGTCGAGGACCTGATGCTGAGAGGGG 6408

Db 6409 TCGGACGGTTCCTTCCACAAATCGACTTGCTCTACGACTACAGACCGTTTTTGGACCA 6468

661 TGTTCGATAGTGCACGATCGGCTTCYTTCCGAGGATGTTCTAAGCCGAAAGTGGCG 720

731 780

Db 6529 GTCATTTGGCGCTGGCATTTCGGACTCGTGTGGCAACGAACCTTCATGCTGGGTA 6588

81 GACGAGGTACAAATAAGAGCGAGGATGCGGTGGAGGCGAGCCTTGGTCACTATG 040

841 TTCAAGGACGCTCCCTAGTGTCTGTGGCCGAATGGGGGCGATCCGATTTCTCTCTGTCGA 900

Db 6649 TTCAGGGACGCTCCTAGTGTCTGTGGCCGAAATGGGGGCGATGCCAATTCTCCTGCTGCA 6708

5768

961 CCCGGCAGTCGACACTTACTTGCTTACCAAGGCGTCCAAATACATGTGAAAGCCGGC 1020

6/69 CCCGGCAGTCGACACTTACCTGGCTACCAAGGCGCTCCATATATATGTCGAAAGCCCGG 8828

Db 6829 CAGCTGCCACCGAAGCTGTTCCATCGCGTTTACAACGGTTGGCGTCCGTTCTTGAAGGAC 6888

1081 GGTTCATGAGCGAGATATGTGTTGGCTTCGCCCTGTCCGTAATTACTCAGGCCTTGAAA 1140

1141 1200

6949 TCAGGACATTAGGTGGCTCATGACTCTGGCAATTTGGCTGAACCGTTTCGGGAGG 7008

QY 1201 GAGTCCTCTCTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGTGTGT-AA 1259  
DB 7009 GAGTCCTCTCTCTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGTGTGTAA 7068  
QY 1260 ACATGAGATTTTCTCTCATGTGAGACTATTCAGACTAATGGAAATAGAGATCTGGGCGG 1319  
DB 7069 ACATGAGATTTTCTCTCATGTGAGACTATTCAGACTAATGGAAATAGAGATCTGGGCGG 7127  
QY 1320 GTTTGGTCAGTTTCTTTTAAAGCGGGTTTATTTAGATCTTCGCTGGTGTATCAACGGATA 1379  
DB 7128 GTTTGGTCAGTTTCTTTTAAAGCGGGTTTATTTAGATCTTCGCTGGTGTATCAACGGATA 7187  
QY 1380 TGAAGAAATACAGCGAGATGTGCCCTGAAGAGATCTCAGAACTTCCAGCTGAGATCCGATC 1439  
DB 7188 TGAAGAAATACAGCGAGATGTGCCCTGAAGAGATCTCAGAACTTCCAGCTGAGATCCGATC 7247  
QY 1440 TGAAGTGTATACGGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGCGAT 1499  
DB 7248 TGAAGTGTATACGGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGCGAT 7307  
QY 1500 TCAGAGGAAAGACAAATAAAGATTAAGAGCTTAAGAGGGGATATCTGAATCTTATGA 1559  
DB 7308 TCAGAGGAAAGACAAATAAAGATTAAGAGCTTAAGAGGGGATATCTGAATCTTATGA 7367  
QY 1560 TAAAGTGTGTGATCAATCTGACTCGCAATATCCAACTGAGGATTCCTGATGATCGA 1619  
DB 7368 TAAAGTGTGTGATCAATCTGACTCGCAATATCCAACTGAGGATTCCTGATGATCGA 7427  
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATC 1679  
DB 7428 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATC 7487  
QY 1680 GTCAAACTCTCTCTGATGATGAGCAAGAAATCTGTTAGACATATCTCTCCGCTTGG 1739  
DB 7488 GTCAAACTCTCTCTGATGATGAGCAAGAAATCTGTTAGACATATCTCTCCGCTTGG 7547  
QY 1740 TGTCTCATGAGAGGAGATCCCAAAAGCAGTATTTGCTGACTATGATGAGAGATCC 1799  
DB 7548 TGTCTCATGAGAGGAGATCCCAAAAGCAGTATTTGCTGACTATGATGAGAGATCC 7607  
QY 1800 GA 1801  
DB 7608 GA 7609

RESULT 3  
E00404  
LOCUS E00404 24595 bp DNA linear PAT 29-SEP-1997  
DEFINITION T1 plasmid T-DNA region.  
ACCESSION E00404.1 GI:2168687  
VERSION E00404.1  
KEYWORDS JP 1985156333-A/1.  
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)  
ORGANISM Agrobacterium tumefaciens  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
REFERENCE 1 (bases 1 to 24595)  
AUTHORS Garti,E.D., Denitsu,D.S. and Richiyado,E.B.  
TITLE SELECTION USING OPINE SYNTHASE GENE  
JOURNAL Patent: JP 1985156333-A 1 16-AUG-1985;  
AGRIJENETRIKUSU RES ASSOC LTD  
COMMENT OS Agrobacterium tumefaciens  
PN JP 1985156333-A/1  
PD 16-AUG-1985  
PR 14-SEP-1984 JP 1984193841  
PI GARTI EE DAARU, DENISU DABURIYU SATSUTON,  
PI RICHIVADO EFU BEIKAA  
PC A01H1/00,C12N5/00,C12N15/00,(C12N5/00,C12R1:91); CC  
strandedness: Double;  
CC topology: Linear;  
CC \*source: strain=15955;  
FH Key Location/Qualifiers

FEATURES  
source Location/Qualifiers  
1..24595  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:358"  
BASE COUNT 6534 a 5509 c 5793 g 6759 t  
ORIGIN  
Query Match 98.0%; Score 1765.2; DB 6; Length 24595;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;  
QY 1 ATGTGAGCTTACCTCTCTGATTAACAGTGGATCATCTCCAACTAAATGTGTGAT 60  
DB 5809 ATGTGAGCTTACCTCTCTGATTAACAGTGGATCATCTCCAACTAAATGTGTGAT 5868  
QY 61 CTGACATGTGTGATTAAGCGGATGATTAAGCCAGGATTCCTGATGATTAAG 120  
DB 5869 CTGACATGTGTGATTAAGCGGATGATTAAGCCAGGATTCCTGATGATTAAG 5928  
QY 121 CGAAGAGCTTCTGAGGAGAGAGATTAATCAAGCTCCACGAGTGCAGCGTGGATTA 180  
DB 5929 CGAAGAGCTTCTGAGGAGAGAGATTAATCAAGCTCCACGAGTGCAGCGTGGATTA 5988  
QY 181 GCTTGCAAAAGGCTGCGCATGTGCTTCCCGAGATCTCAGCTGTGAAAGTAGCA 240  
DB 5989 GCTTGCAAAAGGCTGCGCATGTGCTTCCCGAGATCTCAGCTGTGAAAGTAGCA 6048  
QY 241 GTTCTCTCCGCTTATATCTATATTTGCAAAAGAAATTTCTGGGCGGATCTTGAATGAAA 300  
DB 6049 GTTCTCTCCGCTTATATCTATATTTGCAAAAGAAATTTCTGGGCGGATCTTGAATGAAA 6108  
QY 301 CCTTGGGCGGCGGCAACAGTGTGCTGTTGCCATGAGCTTGGACCATTTTCATG 360  
DB 6109 CCTTGGGCGGCGGCAACAGTGTGCTGTTGCCATGAGCTTGGACCATTTTCATG 6168  
QY 361 GATTTCTCGAAGACAACTAATCCAAAGCCCTGTTTGTGAGCGGTAAAGATGTGA 420  
DB 6169 GATTTCTCGAAGACAACTAATCCAAAGCCCTGTTTGTGAGCGGTAAAGATGTGA 6228  
QY 421 CCGATTGATCTTAGTCATTTTGTGCGCATTTTCAATCTCTAAGATGCGGCTTTCGAAACC 480  
DB 6229 CCGATTGATCTTAGTCATTTTGTGCGCATTTTCAATCTCTAAGATGCGGCTTTCGAAACC 6288  
QY 481 CTGCGAATGCGGCTGATGAGAAATGGCAAGTAAGTAAGTGGTTACCGGTTTACATTAACC 540  
DB 6289 CTGCGAATGCGGCTGATGAGAAATGGCAAGTAAGTAAGTGGTTACCGGTTTACATTAACC 6348  
QY 541 CTGGAAGGCGCGGTGCGCATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 6349 CTGGAAGGCGCGGTGCGCATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 6408  
QY 601 TCGGCAAGTTCCTTTTCAACATGACCTTCTATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 6409 TCGGCAAGTTCCTTTTCAACATGACCTTCTATGATGATGATGATGATGATGATGATGATGATGAT 6468

QY	661	TGTTCCGATAGTGGACGGATATGGCGTCTTTCTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG	720
Db	6469	TGTTCCGATAGTGGACGGATATGGCGTCTTTCTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG	6528
QY	721	GTCAATTGGCGCTGGGCATTTCCGGAATCGTGGTGGCAACGAATCTTTCATGCTGAGGGATA	780
Db	6539	GTCAATTGGCGCTGGGCATTTCCGGAATCGTGGTGGCAACGAATCTTTCATGCTGAGGGATA	6588
QY	781	GACGATGTTACAAATATATGAAAGCAAGTATGCTGTTGGAGGCAAGCTTTGGTCAATGCT	840
Db	6589	GACGATGTTACAAATATATGAAAGCAAGTATGCTGTTGGAGGCAAGCTTTGGTCAATGCT	6648
QY	841	TTCAAGGAGCGCTCTAGTGTGCTGGCCGAATAGGGGGGAGATGGCATTTCCCTCGTGGCA	900
Db	6649	TTCAAGGAGCGCTCTAGTGTGCTGGCCGAATAGGGGGGAGATGGCATTTCCCTCGTGGCA	6708
QY	901	TTTCGCTGTTTCTTCTCTCGAGCGGTATCCGCGCTGCTTTCGATGAGGCGCTTCCCAAT	960
Db	6709	TTTCGCTGTTTCTTCTCTCGAGCGGTATCCGCGCTGCTTTCGATGAGGCGCTTCCCAAT	6768
QY	961	CCCGGACAGTGCACATTAATTGGTCTTACCAAGCGCTCCAAATCATGTGAAAGCCGGG	1020
Db	6769	CCCGGACAGTGCACATTAATTGGTCTTACCAAGCGCGTCCAAATCATGTGAAAGCCGGG	6828
QY	1021	CAGCTGCCACCGAAGCTGTTCCATCCGCTTAAACAACGTTGGCCGCTTCTTAAGGAC	1080
Db	6829	CAGCTGCCACCGAAGCTGTTCCATCCGCTTAAACAACGTTGGCCGCTTCTTAAGGAC	6888
QY	1081	GCTTTTCATGAGCGAGATATTGTGTGGCTTCGCGCTGCTATTAATCTCAGGCGCTTGAA	1140
Db	6889	GCTTTTCATGAGCGAGATATTGTGTGGCTTCGCGCTGCTATTAATCTCAGGCGCTTGAA	6948
QY	1141	TCAGGACACATTAAGTGGGCTCATGACTCTCGCAAAATTTGGCTGAACCGTTTGGGAGG	1200
Db	6949	TCAGGAGACATTAAGTGGGCTCATGACTCTCGCAAAATTTGGCTGAACCGTTTGGGAGG	7008
QY	1201	GAGTCTCTCTCTTCAAGGATAGAGAGATCTTTCTGGGCAACAATCCTCCGTGGTGT-AA	1259
Db	7009	GAGTCTCTCTCTTCAAGGATAGAGAGATCTTTCTGGGCAACAATCCTCCGTGGTGTAA	7068
QY	1260	ACATGGAATTTTCCATGATTTGGGACCTATTCAAGCTTAATGGGAATAGATCTGGCGGG	1319
Db	7069	ACATGGAATTTTCCATGATTTGGGACCTATTCAAGCTTAATGGGAATAGATCTGGCG-GG	7127
QY	1320	GTTTGCTCAGTCTTTTGAAGCGGGTTATTAGATCTCCGCTTGGTGCATCAACGATA	1379
Db	7128	GTTTGCTCAGTCTTTTGAAGCGGGTTATTAGATCTCCGCTTGGTGCATCAACGATA	7187
QY	1380	TGAAGAAATCAGCGGATGTGTCCCTGAAAGGAATCTCAGAACTTCCAGCTGGGATTCGATC	1439
Db	7188	TGAAGAAATCAGCGGATGTGTCCCTGAAAGGAATCTCAGAACTTCCAGCTGGGATTCGATC	7247
QY	1440	TGAAGTGGTTAAACGCTGTGTCTGTGAGCGAGGCAATATGCCATGTTCAAGTCAGGGCGAT	1499
Db	7248	TGAAGTGGTTAAACGCTGTGTCTGTGAGCGAGGCAATATGCCATGTTCAAGTCAGGGCGAT	7307
QY	1500	TCAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAAGACGGGATATCTGAATTTATGA	1559
Db	7308	TCAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAAGACGGGATATCTGAATTTATGA	7367
QY	1560	TAAAGTGGTGTCACTTCGCACTGTGCAAAATATCCAACTCAGGCAATTCCTGCAATCGGA	1619
Db	7368	TAAAGTGGTGTCACTTCGCACTGTGCAAAATATCCAACTCAGGCAATTCCTGCAATCGGA	7427
QY	1620	TACCAATATTTTTCAGGGACAGATGAACAAGCGGTTATTAACGCGCATATGACAGATC	1679
Db	7428	TACCAATATTTTTCAGGGACAGATGAACAAGCGGTTATTAACGCGCATATGACAGATC	7487
QY	1680	GTCAAAATCTTTCCTGATGACTGAAAGCAAAATTCGTGTTAGACCAATATCTCCGCTTGG	1739
Db	7488	GTCAAAATCTTTCCTGATGACTGAAAGCAAAATTCGTGTTAGACCAATATCTCCGCTTGG	7547
QY	1740	TGTTCTCATGAGCGGATTCGAAAGCAAGTATTTGCTCGACATATGAGTGCAGAGATCC	1799

[illegible]



- PUBMED 3021341  
REFERENCE 4 (bases 175278 to 177719)  
AUTHORS Winans,S.C., Allenza,P., Stachel,S.E., McBride,K.E. and Neester,E.W.  
TITLE Characterization of the vire operon of the Agrobacterium Ti plasmid pTiAc
- JOURNAL Nucleic Acids Res. 15 (2), 825-837 (1987)  
MEDLINE 87146396  
PUBMED 3547330  
5 (bases 50125 to 54471)  
AUTHORS Habeeb,L.F., Wang,L. and Winans,S.C.  
TITLE Transcription of the octopine catabolism operon of the Agrobacterium tumor-inducing plasmid pTiAc is activated by a tyler-type regulatory protein
- JOURNAL Mol. Plant Microbe Interact. 4 (4), 379-385 (1991)  
MEDLINE 92190631  
PUBMED 17996699  
6 (bases 50125 to 54471)  
AUTHORS Valdivia,R.H., Wang,L. and Winans,S.C.  
TITLE Characterization of a putative periplasmic transport system for octopine accumulation encoded by Agrobacterium tumefaciens Ti plasmid pTiAc
- JOURNAL J. Bacteriol. 173 (20), 6398-6405 (1991)  
MEDLINE 92011411  
PUBMED 1655707  
7 (bases 68024 to 76460)  
AUTHORS Kim,K.S. and Farrand,S.K.  
TITLE Ti plasmid-encoded genes responsible for catabolism of the crown gall opine mannopine by Agrobacterium tumefaciens are homologs of the T-region genes responsible for synthesis of this opine by the plant tumor
- JOURNAL J. Bacteriol. 178 (11), 3275-3284 (1996)  
MEDLINE 96236046  
PUBMED 8655509  
8 (bases 39167 to 45466)  
AUTHORS Fuqua,C. and Winans,S.C.  
TITLE Localization of OccR-activated and Trar-activated promoters that express two ABC-type permeases and the trar gene of Ti plasmid pTiR10
- JOURNAL Mol. Microbiol. 20 (6), 1199-1210 (1996)  
MEDLINE 96405643  
PUBMED 8809772  
9 (bases 22855 to 39243; 96353 to 111409)  
AUTHORS Alt-Worbe,J., Styrker,J.L., Fuqua,C., Li,P.L., Farrand,S.K. and Winans,S.C.  
TITLE The conjugal transfer system of Agrobacterium tumefaciens octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes
- JOURNAL J. Bacteriol. 178 (14), 4248-4257 (1996)  
MEDLINE 96312368  
PUBMED 8763954  
10 (bases 54312 to 62806)  
AUTHORS Oger,P., Kim,K.S., Sackett,R.L., Piper,K.R. and Farrand,S.K.  
TITLE Octopine-type Ti plasmid code for a mannopine-inducible dominant-negative allele of trar, the quorum-sensing activator that regulates Ti plasmid conjugal transfer
- JOURNAL Mol. Microbiol. 27 (2), 277-288 (1998)  
MEDLINE 98143415  
PUBMED 9484884  
11 (bases 133963 to 138972)  
AUTHORS Kaloogeraki,V.S. and Winans,S.C.  
TITLE Wound-released chemical signals may elicit multiple responses from an Agrobacterium tumefaciens strain containing an octopine-type Ti plasmid
- JOURNAL J. Bacteriol. 180 (21), 5660-5667 (1998)  
MEDLINE 99009000  
PUBMED 9791116  
12 (bases 76455 to 98723)  
AUTHORS Lyi,S.M., Jafari,S. and Winans,S.C.  
TITLE Mannopinic acid and agropinic acid catabolism region of the octopine-type Ti plasmid pTi15955
- JOURNAL Mol. Microbiol. 31 (1), 339-347 (1999)  
MEDLINE 99141607  
PUBMED 9987134
- REFERENCE 13 (bases 1 to 194140)  
AUTHORS Winans,S.C., Zhu,J., Oger,P.M., Schrammeijer,B., Hooykaas,P.J. and Farrand,S.K.  
TITLE Octopine-type Ti plasmid sequence unpublished
- JOURNAL 14 (bases 68024 to 76460)  
AUTHORS Kim,K.-S.  
TITLE Direct Submission
- JOURNAL Submitted (10-JAN-1995) Kun-Soo Kim, Microbiology, University of Illinois at Urbana-Champaign, 1201 W. Gregory Dr., Urbana, IL 61801, USA
- REFERENCE 15 (bases 96353 to 111409)  
AUTHORS Winans,S.C.  
TITLE Direct Submission
- JOURNAL Submitted (20-DEC-1995) Stephen C. Winans, Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 16 (bases 39167 to 45466)  
AUTHORS Fuqua,C.  
TITLE Direct Submission
- JOURNAL Submitted (07-FEB-1996) Clay Fuqua, Biology, Trinity University, 715 Stadium Dr., San Antonio, TX 78212, USA
- REFERENCE 17 (bases 54312 to 62806)  
AUTHORS Kim,K.-S., Sackett,R.L. and Farrand,S.K.  
TITLE Direct Submission
- JOURNAL Submitted (04-JUN-1996) Crop Sciences, University of Illinois at Urbana-Champaign, 240 ERMU, 1201 W. Gregory Dr., Urbana, IL 61801, USA
- REFERENCE 18 (bases 186239 to 193239)  
AUTHORS Zhu,J. and Winans,S.C.  
TITLE Direct Submission
- JOURNAL Submitted (17-NOV-1997) Section of Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 19 (bases 76455 to 98723)  
AUTHORS Winans,S.C. and Lyi,S.  
TITLE Direct Submission
- JOURNAL Submitted (20-NOV-1997) Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 20 (bases 170209 to 175283)  
AUTHORS Zhu,J. and Winans,S.C.  
TITLE Direct Submission
- JOURNAL Submitted (25-NOV-1997) Section of Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 21 (bases 133963 to 138972; 176911 to 186244)  
AUTHORS Styrker,J.L., Mantis,N.J., Kaloogeraki,V.S. and Winans,S.C.  
TITLE Direct Submission
- JOURNAL Submitted (23-DEC-1997) Microbiology, Cornell University, Ithaca, NY 14853, USA
- REFERENCE 22 (bases 22855 to 39243)  
AUTHORS Winans,S.C.  
TITLE Direct Submission
- JOURNAL Submitted (03-AUG-1998) Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 23 (bases 62801 to 66985)  
AUTHORS Oger,P.M.  
TITLE Direct Submission
- JOURNAL Submitted (29-DEC-1998) Crop Sciences, University of Illinois at Urbana-Champaign, 361 ERMU, 1201 W. Gregory Dr., Urbana, IL 61801, USA
- REFERENCE 24 (bases 110822 to 133968)  
AUTHORS Zhu,J. and Winans,S.C.  
TITLE Direct Submission
- JOURNAL Submitted (18-AUG-1999) Section of Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE 25 (bases 1 to 194140)  
AUTHORS Zhu,J., Oger,P.M., Schrammeijer,B., Hooykaas,P.J., Farrand,S.K. and Winans,S.C.  
TITLE Direct Submission
- JOURNAL Submitted (07-MAR-2000) Microbiology, Cornell University, Wing Hall, Ithaca, NY 14853, USA
- COMMENT On or before Jun 20, 2000 this sequence version replaced  
gi:3377769, gi:1215729, gi:1381799, gi:797330, gi:2665704,  
gi:2749889, gi:5758260, gi:2773254, gi:2773255, gi:2773256,  
gi:154820, gi:154827, gi:2662524, gi:39122, gi:2687604, gi:154767,

```
FEATURES
    gi:154773, gi:6689862.
    Location/Qualifiers
        source
            /organism="Agrobacterium tumefaciens"
            /mol_type="genomic DNA"
            /db_xref="taxon:358"
            /plasmid="T1"
            /note="extrachromosomal
            octopline-lyse"
            13244..13618
        source
            /organism="Agrobacterium tumefaciens"
            /mol_type="genomic DNA"
            /db_xref="taxon:358"
            /insertion_seq="ybd"
            13636..14181
        source
            /organism="Agrobacterium tumefaciens"
            /mol_type="genomic DNA"
            /db_xref="taxon:358"
            /insertion_seq="ybe"
            /note="similar to 1866"
            14183..14666
        source
            /organism="Agrobacterium tumefaciens"

Query Match      98.0%; Score 1765.2; DB 1; Length 194140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY      1  ATGTGACCTTCACTCTCTCTTGAATACCAAGTGCATATCTCCACCAAAATGTGGAT 60
DB      4902 ATGTGACCTTCACTCTCTTGTATACCAAGTGCATATCTCCACCAAAATGTGGAT 4961
QY      61  CTGACATAGTGTGATAGGGGATGATGATGACCCGAGGTTCCGATGCTTTAGAA 120
DB      4962 CTGACATAGTGTGATAGGGGATGATGATGACCCGAGGTTCCGATGCTTTAGAA 5021
QY      4962 CTGACATAGTGTGATAGGGGATGATGATGACCCGAGGTTCCGATGCTTTAGAA 5021
DB      121  CGAAGACTTCTAGGGGAGAGGATTACTCAAGCTCCACCGAGTGCAGCGTGGTTA 180
QY      5022 CGAAGACTTCTAGGGGAGAGGATTACTCAAGCTCCACCGAGTGCAGCGTGGTTA 5081
DB      181  GCTTGCAGAAAGCTGCGCATGTGCTGCCGAGATCTCAGCTGTGAAAGGTGCA 240
QY      5082 GCTTGCAGAAAGCTGCGCATGTGCTGCCGAGATCTCAGCTGTGAAAGGTGCA 5141
DB      241  GTTCTCCGCTTATATCTATATTGGCAAGAAATCTGGGGGCGGATCTGAATGAAA 300
QY      5142 GTTCTCCGCTTATATCTATATTGGCAAGAAATCTGGGGGCGGATCTGAATGAAA 5201
DB      301  CCTTGGGCGGGGCAACAGTAGTGTCTGTTGCCATCGACTTGGACACCATTTGCATG 360
QY      5202 CCTTGGGCGGGGCAACAGTAGTGTCTGTTGCCATCGACTTGGACACCATTTGCATG 5261
DB      361  GATTTCCTCGAAGACAACTAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGTGTGA 420
QY      5262 GATTTCCTCGAAGACAACTAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGTGTGA 5321
DB      421  CGCATGATCTTGTGCAATTCGTCGTCATTTCAATCTCTAAGACTGCCGCTTTTCAAC 480
QY      5322 CGCATGATCTTGTGCAATTCGTCGTCATTTCAATCTCTAAGACTGCCGCTTTTCAAC 5381
DB      481  CTGCCAATGCCGCTGTACGAGAAATGCGACATGAAATGCGTTACCGGGTTTACATACC 540
QY      5382 CTGCCAATGCCGCTGTACGAGAAATGCGACATGAAATGCGGGTTTACATACC 5441
DB      541  CTGGAAGGGGCGGTCATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY      5442 CTGGAAGGGGCGGTCATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5501
DB      601  TCGGAGAGTCTCTTCCAAATGCAATGCTCTGATGATGATGATGATGATGATGATGATGAT 660
QY      5502 TCGGAGAGTCTCTTCCAAATGCAATGCTCTGATGATGATGATGATGATGATGATGATGAT 5561
DB      661  TGTTCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
```

```
DB      5562 TGTTCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5621
QY      721  GTGATTTGGGCTGGGATTTCCGGAGCTGTTGGGCAAGAACTGCTTCAATGCTGGGGTA 780
DB      5622 GTGATTTGGGCTGGGATTTCCGGAGCTGTTGGGCAAGAACTGCTTCAATGCTGGGGTA 5681
QY      781  GACATGTTTACATATATATGAGCAAGTGTATGCTGTGAGGCAAGCTTTGTCATATGCT 840
DB      5682 GACATGTTTACATATATATGAGCAAGTGTATGCTGTGAGGCAAGCTTTGTCATATGCT 5741
QY      841  TTCAGGAGCGCTCTATAGTGTGTCGCGCAATGAGGCGCATGCGATTTCTCTCTGCA 900
DB      5742 TTCAGGAGCGCTCTATAGTGTGTCGCGCAATGAGGCGCATGCGATTTCTCTCTGCA 5801
QY      901  TTCTGCTGTTTTTTCTTCTCGAGCGCTTACGCGCTCTCTGATGAGCGCTTCCAAAT 960
DB      5802 TTCTGCTGTTTTTTCTTCTCGAGCGCTTACGCGCTCTCTGATGAGCGCTTCCAAAT 5861
QY      961  CCGGACACAGTGCACACTTACTTGTCTTACCAAGCGTCCAAATACATGTGAAAGCCGG 1020
DB      5862 CCGGACACAGTGCACACTTACTTGTCTTACCAAGCGTCCAAATACATGTGAAAGCCGG 5921
QY      1021 CAGCTGCCACCGAAGCTGTTCCATGCGCTTTTCAACGCGTTGGCGCTTTTGAAGAC 1080
DB      5922 CAGCTGCCACCGAAGCTGTTCCATGCGCTTTTCAACGCGTTGGCGCTTTTGAAGAC 5981
QY      1081 GGTTCCTCATGAGGAGATATGTTGTTGGCTTCCCTCTGCTTATTAATCAGGCTTGAAA 1140
DB      5982 GGTTCCTCATGAGGAGATATGTTGTTGGCTTCCCTCTGCTTATTAATCAGGCTTGAAA 6041
QY      1141 TCAGACACATTAGTGGGCTCATGACTCTGCAAAATTTGGTGAACGTTTGGGAGG 1200
DB      6042 TCAGACACATTAGTGGGCTCATGACTCTGCAAAATTTGGTGAACGTTTGGGAGG 6101
QY      1201 GAGTCTTCTTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTCTGTGTGT-AA 1259
DB      6102 TCGAGACATTATAGTGGGCTCATGACTCTGCAAAATTTGGTGAACGTTTGGGAGG 6161
QY      1260 ACATGAGTTTTCTCTATGATTTGGGACCTTATTCAGCTAATGGGAAATAGGATTTGGCGG 1319
DB      6162 ACATGAGTTTTCTCTATGATTTGGGACCTTATTCAGCTAATGGGAAATAGGATTTGGCGG 6220
QY      1320 GTTGTGCTGCTTTTGAAGCGGGTTTATGAGATCTCTCGCTGTGATCAACGGATA 1379
DB      6221 GTTGTGCTGCTTTTGAAGCGGGTTTATGAGATCTCTCGCTGTGATCAACGGATA 6280
QY      1380 TGAAGAAATTCAGCGGATGTGCTTGAAGAAATCTCAGAACTTCCAGTTCGATC 1439
DB      6281 TGAAGAAATTCAGCGGATGTGCTTGAAGAAATCTCAGAACTTCCAGTTCGATC 6340
QY      1440 TGAAGTGTTAACGGTGTGTCTGTAGCCAGCGCATATGCTATGTTCAAGTCAAGGCGAT 1499
DB      6341 TGAAGTGTTAACGGTGTGTCTGTAGCCAGCGCATATGCTATGTTCAAGTCAAGGCGAT 6400
QY      1500 TCAGAGAAAAAGACAAAATATAAGATAAGGCTTAAGAGCGGGATCTGAACTTTATGA 1559
DB      6401 TCAGAGAAAAAGACAAAATATAAGATAAGGCTTAAGAGCGGGATCTGAACTTTATGA 6460
QY      1560 TAAGTGTGTGATCATTGCACTCGCAATATCCAACTCAGGACTTGCCTGATGCGA 1619
DB      6461 TAAGTGTGTGATCATTGCACTCGCAATATCCAACTCAGGACTTGCCTGATGCGA 6520
QY      1620 TACCAATATTTTTCAGGACCACTGAAACCAAGCGGTTGATTAACAGCCATATGACAGATC 1679
DB      6521 TACCAATATTTTTCAGGACCACTGAAACCAAGCGGTTGATTAACAGCCATATGACAGATC 6580
QY      1680 GTCAAAATCTTCTCTGATGATGAAAGAAATTTCTGTTAGACATATCTCCGCTTG 1739
DB      6581 GTCAAAATCTTCTCTGATGATGAAAGAAATTTCTGTTAGACATATCTCCGCTTG 6640
QY      1740 TGTCTCATGAGACGGGATGCAAAAGCAAGTGTATGCTCTGACTATGATGCTGAGATCC 1799
DB      6641 TGTCTCATGAGACGGGATGCAAAAGCAAGTGTATGCTCTGACTATGATGCTGAGATCC 6700
```

QY 1800 GA 1801  
Db 6701 GA 6702

RESULT 6  
TIMPMS2  
LOCUS  
DEFINITION TIMPMS2 2591 bp DNA linear BCT 19-MAR-1996  
Ti plasmid (A. tumefaciens octopline strain) tms1 gene, complete

ACCESSION K02554  
VERSION K02554.1 GI:154747

KEYWORDS  
SOURCE  
ORGANISM  
plasmid Ti

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

1 (bases 1 to 2591)  
Klee, H., Montoya, A., Horodyski, F., Lichtenstein, C., Garfinkel, D.,  
Fuller, S., Flores, C., Peschon, J., Nester, E. and Gordon, M.  
Nucleotide sequence of the tms gene of the pTiA6NC octopline Ti  
plasmid: two gene products involved in plant tumorigenesis  
Proc. Natl. Acad. Sci. U.S.A. 81 (6), 1728-1732 (1984)  
84170374  
6584906

Original source text: Plasmid Ti (clone: pTiA6NC.).  
The tms1 protein has a region which showed significant homology to  
the adenine binding region of p-hydroxybenzoate hydroxylase from  
P. fluorescens. It seems that the tms1 protein binds adenine  
either as substrate or cofactor.  
The initiation and termination sites of transcription have not yet  
been identified [1].

FEATURES  
source  
Location/Qualifiers  
1..2591  
/organism="Plasmid Ti"  
/mol\_type="genomic DNA"  
/specific\_host="octopline"  
/db\_xref="taxon:2512"  
/clone="pTiA6NC."  
/plasmid="Plasmid Ti"  
160..2427  
/gene="tms1"  
160..2427  
/gene="tms1"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AAA92550.1"  
/db\_xref="GI:154750"  
/translation="MSASPLLDNQCCHLPTKMYDLTWVDKADELDKRVSDAPLERNS  
RGRITQISTECAGLACKRLADGRPEISAGKVAIVASVYIGKEILGRILSKPM  
ARATVGLVAILDAPFCMDRSEALQIALFLSGRCAPLIDLPVAILSTKRAFPRT  
LNPVLVNGTKMCTVGTFTTLEBAGVPPDMVAYGRNMLKKSASSPTIIDLYRPF  
DOCSDSGRIGFPEVDPKPVAVYAGISLVANELLHAGVDVITYEASDVGGFL  
WSHAFRAPSVAEMGMRPPAPAFCLFFELERYGLSSMPFNPGETVDYLYVQVQ  
YMKAKGLPKLPHRYVNGWRAFLKGFYRDLVASPVATIALKSGDFRMAHDSQ  
ITLNRFGRESFGSIEIRIFLGTHPGEETSPFDMDLFLMGIGSGGFGFVPSGFI  
EILRLVINGEENORMCEGISLPRILAEVNGVSVSORICHVOYRAIOKXKTKIX  
IRKSGISELYDYVYVTSGLANTOLRCLICDTRNIPAPVNOAVNDSHMTGSKPLM  
TERKFWLDHILPSCLMDGIAKAVYCLDYEPDOPNGGLVLIYTWEDSHKLLAVPD  
KKEKRLCLRDALISRSFPAPFOHLFPACADYDQVITQDMVLTIDENAGAFILNRGDF  
YSEELFOALDITANDITGVYLAGCSGFTGVEGAIGTACNAVCAIITHNGGILAKN  
PLEHSMKRYVYRFRN"

BASE COUNT 654 a 570 c 656 g 711 t  
ORIGIN About 250 bp after segment 1.

Query Match 97.7%; Score 1760.4; DB 1; Length 2591;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1789; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 1 ATGTCACTTCACTCTCTCTTGATACCAAGTCGATCATCTCCCAACCAAAATGGTGAT 60  
Db 160 ATGTCACTTCACTCTCTTGATACCAAGTCGATCATCTCCCAACCAAAATGGTGAT 219

QY 61 CTGACATGTCGATTAAGCGCGATGAATTGSAACGAGGGTTTCCGATGCTTCTTAGAA 120  
Db 220 CTGACATGTCGATTAAGCGCGATGAATTGSAACGAGGGTTTCCGATGCTTCTTAGAA 279

QY 121 CGAAGACTTCTTAGGGGAAAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGGTTA 180  
Db 280 CGAAGACTTCTTAGGGGAAAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGGTTA 339

QY 181 GCTTGAAAAGGTGGCCGATGTGCTCCCGAGATCTCAGCTGTGGAAGTAGCA 240  
Db 340 GCTTGAAAAGGTGGCCGATGTGCTCCCGAGATCTCAGCTGTGGAAGTAGCA 399

QY 241 GTTCTCTCGCTTATATCTATATTTGGCAAGAAATCTGGGCGGATCTTGAATCGAAA 300  
Db 400 GTTCTCTCGCTTATATCTATATTTGGCAAGAAATCTGGGCGGATCTTGAATCGAAA 459

QY 301 CTTTGGGCGGCGCAACAGTAGTGTCTCTGTTCCATCGACTTGGACACCATTTGCATG 360  
Db 460 CTTTGGGCGGCGCAACAGTAGTGTCTCTGTTCCATCGACTTGGACACCATTTGCATG 519

QY 361 GATTTCTCGAAGCAACATCAATCCAGCCGTTTGTGCGGGTAAAGATGCA 420  
Db 520 GATTTCTCGAAGCAACATCAATCCAGCCGTTTGTGCGGGTAAAGATGCA 579

QY 421 CCGATTGATCTTATGATCTTCTGCGCATTTCAATCTTAAGACTCGCGCTTTCGAACC 480  
Db 580 CCGATTGATCTTATGATCTTCTGCGCATTTCAATCTTAAAGACTCGCGCTTTCGAACC 639

QY 481 CTGCGAATGCGGCTGTACAGAAATGGCAGATGAATGCGTTACCGGTTTACATTAAC 540  
Db 640 CTGCGAATGCGGCTGTACAGAAATGGCAGATGAATGCGTTACCGGTTTACATTAAC 699

QY 541 CTTGAAGGGGCGGTGCGATTTGACATGTAGCTTATAGTGTGAACCTGATGTGAAGGT 600  
Db 700 CTTGAAGGGGCGGTGCGATTTGACATGTAGCTTATAGTGTGAACCTGATGTGAAGGT 759

QY 601 TCGGCAAGTTCTCTTCAACAATCGACTTCTACAGCTGACAGCCGTTTTTGAACAA 660  
Db 760 TCGGCAAGTTCTCTTCAACAATCGACTTCTACAGCTGACAGCCGTTTTTGAACAA 819

QY 661 TGTTCGATATGTGACGAGTCGCTCTTCTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 720  
Db 820 TGTTCGATATGTGACGAGTCGCTCTTCTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 879

QY 721 GTCAATGGGCGTGGCATTTCCGGACTCGTGTGGCAACGAATGCTCATGCTGGGGTA 780  
Db 880 GTCAATGGGCGTGGCATTTCCGGACTCGTGTGGCAACGAATGCTCATGCTGGGGTA 939

QY 781 GACGATGTTACATATATGAGCAAGTGAATCGTGTGAGGCAAGCTTTGTGACATGCT 840  
Db 940 GACGATGTTACATATATGAGCAAGTGAATCGTGTGAGGCAAGCTTTGTGACATGCT 999

QY 841 TTAGGAGAGCTCTTGTGTGTGTGCGCGCAAAATGGGGCCATGCGATTTCTCTGCTGCA 900  
Db 1000 TTAGGAGAGCTCTTGTGTGTGTGCGCGCAAAATGGGGCCATGCGATTTCTCTGCTGCA 1059

QY 901 TTCTGCTGTTTTTCTTCCCTGAGCGTTACGGCTGTTCTGATGAGCGCGTCCCAAT 960  
Db 1060 TTCTGCTGTTTTTCTTCCCTGAGCGTTACGGCTGTTCTGATGAGCGCGTCCCAAT 1119

QY 961 CCGGCAAGTGCACACTTACTTGTCTACCAAGGCGTCCCAATCATGTGGAAGCGGG 1020  
Db 1120 CCGGCAAGTGCACACTTACTTGTCTACCAAGGCGTCCCAATCATGTGGAAGCGGG 1179

QY 1021 CAGCTCCACCGAAGCTGTTCCATCGCGTTTACACAGTGTGGCGTCTTGAAGGAC 1080  
Db 1180 CAGCTCCACCGAAGCTGTTCCATCGCGTTTACACAGTGTGGCGTCTTGAAGGAC 1239

QY 1081 GGTTCATGAGGAGATATGTGTTGGCTGCGCTGCTGCTTCTCAGGCGCTGAAA 1140  
Db 1240 GGTTCATGAGGAGATATGTGTTGGCTGCGCTGCTGCTTCTCAGGCGCTGAAA 1299

QY	1141	TCAGAGACATTAGGTGGGCTCATGACCTCCGGGCAAAATTTGGCTGAACCGTTTGGGAGG	1200
Db	1300	TCAGAGACATTAGGTGGGCTCATGACCTCCGGGCAAAATTTGGCTGAACCGTTTGGGAGG	1359
QY	1201	GAGTCTCTCTTTCAGGGATAGAGAGATCTTTTCGGGACACATCCCTCGTGGT-AA	1259
Db	1360	GAGTCTCTCTTTCAGGGATAGAGAGATCTTTTCGGGACACATCCCTCGTGGTGA	1419
QY	1260	ACATGAGATTTTCCCTCATGATTTGGGACCTTAACTTAATGGGAATAGATCTGGCGGG	1319
Db	1420	ACATGAGATTTTCCCTCATGATTTGGGACCTTAACTTAATGGGAATAGATCTGGC-GG	1478
QY	1320	GTTCGCTCAGTTTGTGAAGGGGGTTTATTAGATCTCTCCGCTTGGTCATCAACGATA	1379
Db	1479	GTTCGCTCAGTTTGTGAAGGGGGTTTATTAGATCTCTCCGCTTGGTCATCAACGATA	1538
QY	1380	TGAAGAAATCAGCGGATGTGCCCTGAAGAAATCTCAGAACCTTCCAGCTCGGATCGCATC	1439
Db	1539	TGAAGAAATCAGCGGATGTGCCCTGAAGAAATCTCAGAACCTTCCAGCTCGGATCGCATC	1598
QY	1440	TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGCGAT	1499
Db	1599	TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGCGAT	1658
QY	1500	TCAGAGAGAAAGAACAAAATTAAGATTAAGCTTAAAGCCGGATATTTGAACCTTTATGA	1559
Db	1659	TCAGAGAGAGAAAGAACAAAATTAAGATTAAGCTTAAAGCCGGATATTTGAACCTTTATGA	1718
QY	1560	TAAAGTGTGTGCATCATCTGGACTGGCAAAATATCCAACTCAGAGGCAATGCCATGATGGA	1619
Db	1719	TAAAGTGTGTGCATCATCTGGACTGGCAAAATATCCAACTCAGAGGCAATGCCATGATGGA	1778
QY	1620	TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACAGCCATATAGACGATC	1679
Db	1779	TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACAGCCATATAGACGATC	1838
QY	1680	GTCAAAATCTTCTCTGATGACTGAACGAAAATTTCTGTTAACAATATCTCCGCTCTTG	1739
Db	1839	GTCAAAATCTTCTCTGATGACTGAACGAAAATTTCTGTTAACAATATCTCCGCTCTTG	1898
QY	1740	TGTCCTCATGACCGGATCGCAAAAGCAAGTATTCCTGACATATAGTGTGACGAGATCC	1799
Db	1899	TGTCCTCATGACCGGATCGCAAAAGCAAGTATTCCTGACATATAGTGTGACGAGATCC	1958
QY	1800	GA 1801	
Db	1959	GA 1960	
RESULT 7			
AB016260	206479 bp	DNA	circular BCT 12-JUN-2003
LOCUS	AB016260		
DEFINITION	Agrobacterium tumefaciens plasmid pTi-SAKURA, complete sequence.		
ACCESSION	AB016260	AB007145	AB007188
VERSION	AB016260.1	GI:6498173	
KEYWORDS			
SOURCE			
ORGANISM	Agrobacterium tumefaciens (Rhizobium radiobacter)		
REFERENCE			
AUTHORS	Hattori,Y., Suzuki,K., Uraji,M., Ohta,N., Katoh,A. and Yoshida,K.		
TITLE	Genome structure of pTi-SAKURA (1): strategy for DNA sequencing of a Japanese cherry-Ti plasmid		
JOURNAL	Nucleic Acids Symp. Ser. 37, 159-160		(1997)
MEDLINE	98247140		
PUBMED	9586048		
REFERENCE			
AUTHORS	Hattori,Y., Suzuki,K., Ohta,N., Uraji,M., Katoh,A. and Yoshida,K.		
TITLE	Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens		
JOURNAL	Nucleic Acids Symp. Ser. 39, 265-266		(1998)
REFERENCE			
	3		

FEATURES	source
COMMENT	g1:2962699, g1:2982703. Location/Qualifiers
FEATURES	1. .206479 /organism="Agrobacterium tumefaciens" /mol_type="genomic DNA" /strain="MAF301001" /db_xref="taxon:358" /plasmid="pTi-SAKURA" /note="collection site: Japan (from cherry tree)" join(205959. .206479,1. .1435) /gene="t:lorf195" join(205959. .206479,1. .1435) /gene="t:lorf195" /note="hyub gene homolog" /codon_start=1 /transl_table=11 /protein_id="BAA87820.1" /db_xref="GI:6498368" /translation="MDAVRTVAVNNRNFAIVEEASAIYRTAHTTFFKIVODYOCAIA TAGGEMAYPMLGCVNVFVGSPLKPTLDAIGRNLRKGDIIITNDPATGLYTHLMD VTLLEPIFDEKLIAGMAFVHSDIGAVPGSISPATFVFOEGLRVRPMKLYEGGV LNAIKSIFODNSRIPTELMGDIOMISGLKSMDRVSELCEBRYDRQSVTEGMODVIN YAKTKAVARTRTIDGVYTRTDYLTBEGHDOQLAVFSYTWVKQSEIEVDPTGDPOLA AAINLVGATHTPRTIYTOCLYAVILTDVDTLPNSGJIRALHGAIPRTVINAAYPASG GSSAASATRAYDIVILGCLNQLPEGLAAGGSGVIVASAPDPRGRDNNVNGTID GGGARGVGDGLGSEVRYGORSVPAYVEIIEVLWMRALPDPDSRARGFASGAL EIEEMTNSRAVITVNRINLFVAPMGFKCEIIGLCKAVNPGRPDSRGVKSIVLE LGGDILIRITSSTVGARPELNRDVAAIEAIENGMISPERADYAVAVPDRSKIDE AAIAREHRERANKSDDFNFCYRQRODQWSQOTRELASRALTEQRIRSQLVDRV HHRLLAGERVNDAGKLDQVIAEAGRL" complement(1559. .2479) /gene="t:lorf1" complement(1559. .2479) /gene="t:lorf1" /note="Putative transcriptional regulatory gene" /codon_start=1 /transl_table=11 /protein_id="BAA87626.1" /db_xref="GI:6498174" /translation="MSYQVIEIGSRGVVAAEHLFRRAIRLMSOPTLSOQIKRLE ELMGVQLFTRSTRVVDLTRGASFPFHARDILKIDBAIINAKLEAGGSGPGEELII GAISPEQLPLISRFORPRTLAIKEVDSIELRGIERGDPRHVGMPNSN LVFPHLSLRFVAVMPROSAALKARLRLSDPAGVGFALKEFELSCEFTYVEOLSG
REFERENCE	4 Uraji,M., Suzuki,K., Ohta,N., Hattori,Y., Kato,A. and Yoshida,K. Genome structure of pTi-SAKURA (TII): Characteristics of T-DNA Nucleic Acids Symp. Ser. 39, 185-186 (1998)
REFERENCE	5 Suzuki,K., Ohta,N., Hattori,Y., Uraji,M., Kato,A. and Yoshida,K. Novel structural difference between nopaline- and octopine-type trp genes: construction of genetic and physical map and sequencing of trp/trai and rep gene clusters of a new Ti plasmid pTi-SAKURA Biochim. Biophys. Acta 1396 (1), 1-7 (1998)
REFERENCE	6 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	7 Suzuki,K. and Yoshida,K. Direct Submission Submitted (22-JUL-1998) Kazuo Yoshida, Hiroshima University, Faculty of Science, Kagamiyama 1-3-1, Higashihiroshima-shi, Hiroshima 739, Japan (E-mail:k-suzuki@ipc.hiroshima-u.ac.jp, Tel:0824-24-7456, Fax:0824-24-7033) On or before May 14, 2001 this sequence version replaced g1:2962699, g1:2982703. Location/Qualifiers
REFERENCE	8 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	9 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	10 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	11 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	12 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	13 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	14 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	15 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	16 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	17 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	18 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	19 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	20 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	21 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	22 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	23 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	24 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	25 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	26 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	27 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	28 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	29 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	30 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	31 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	32 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)
REFERENCE	33 Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K., Kato,A. and Yoshida,K. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid Gene 242 (1-2), 331-336 (2000)

AGLIENDIIVSDTNTALATVAGVGTIFLPEWVSAMAPDGVTVRVVDMSVEISMAY  
GMPDPCVPGLIFPEVEFATLAONLISARK"  
/complement (2594. .3319)  
/gene="tlorf2"  
/complement (2594. .3319)  
/gene="tlorf2"  
/function="involved in nopaline transport"  
/note="Probable nocm gene"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="BAA87627.1"  
/db\_xref="GI:6498175"  
/translation="MDIQLIESEPKLLAAVPTTLTAFISLIGFVVSYPVALMRLS  
KNRIVSLAYGVYIRSTPLVOMFLIYGAQFQVLSVGLMSFREPFCALIA  
LAINTAYTSEIRGSIQSVLSIQIEBARAVGMSFLOPRIVIPILAIROLPAYGNE  
VMILKSTSLASTITIVETGLAKOITISATYSPVEVITVAGAIYLFITFVVSRLVMLA  
EMWLNPMRVARVGGTAPKAEATH"  
complement (3321. .4031)  
/gene="tlorf3"  
complement (3321. .4031)  
/gene="tlorf3"  
/function="involved in nopaline transport"  
/note="Probable nocm gene"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="BAA87628.1"  
/db\_xref="GI:6498176"  
/translation="MDITLQMGDAGCDELARGAMNTVVVAACTYFPGIIFGSLPRA  
AKLSRFWSLRLDGVYTVVGVPELITLVFPGGSLTKRTIANGLFEGYIEPT  
FVIGVLCISVAGVATEVIRAAVLAVPQOIEAKSIGMPWLRRLRVILPOAARA  
LPGIAGVMOFTLKDTSLSIVGVEIWRITAMGAGSTKQPFYITAFVFLILSSVS  
NRGFLKAEKXANGVRSQ"  
complement (4106. .4957)  
/gene="tlorf4"  
complement (4106. .4957)  
/gene="tlorf4"  
/function="involved in nopaline transport  
(nopaline-binding protein)"  
/note="Probable nocm gene"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="BAA87629.1"  
/db\_xref="GI:6498177"  
/translation="MKFENLALAAVTVGLAAGPTQAKDVYSITATEGSYAPNF  
KDAGKLGIPDLDGNDLCRNMIECKFEVQAMDGILPSLTAAGYDAIMAAKIOPAR  
BKTIARSPRYLTPMTPLTTADSELTQVAIRNLPIDNTTPQKALDKPTIFBEV  
KFGVQAGTSHAFPKQMPSPVQISTYDTIDNVMDLAKRIDSLASVSLKEPLTDRP  
DNDDLKMFGRMTGGLFGKGVGIRKEDALRALFDKALDAADGTQKLSQWFG  
YDASPKQ"  
complement (5022. .5795)  
/gene="tlorf5"  
complement (5022. .5795)  
/gene="tlorf5"  
/function="involved in nopaline transport"  
/note="Probable nocm gene"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="BAA87630.1"  
/db\_xref="GI:6498178"  
/translation="MDATQPTLVADVHKNFGLTEVLKGISLTANKGDVVSITGSGS  
GKSTFLCANFLETPNKRIAVQEEVVVTTDAAGRLIGVDRKKIERMQLCMVFS  
FNIMGHMTVLQNWEGFLVKKOPKGEVRADMFLLKGIANKHAAVPSOLGSGQOO  
RVSITARLAMPQALLPDEPTSLDPELVSEVVKTRKLABEGRTWVVVTHMGFPAD  
VSSKVLFEKQIETGCTPOEVFONPISPCRAFLSSVL"  
6016. .6951  
/gene="tlorf6"  
6016. .6951  
/gene="tlorf6"  
/function="transcriptional regulator for nopaline  
catabolism"  
/note="Probable nocm gene"  
/codon\_start=1  
/transl\_table=1

/protein\_id="BAA87631.1"  
/db\_xref="GI:6498179"  
/translation="MRPCVFMKEKAMIOSROLEAPRAVWLGTGMSANLVRITOPAI  
SRIRLDEEIGSLFERGONRLPRREGILFEKYSRFPNGIOHDXAALRKSHM  
GSLRVACTYAPALSWFSWGVITQTIADRPVSVYLTDPVQSTVLEISLVQHDGISIL  
AGDYPALTTEPVSPFVAVCLLPGRHLEKTEVHAUDELGESLILCSFVSLMQDTDA  
ALDSCVHNCNRRIEESLALNCLDLSVRGSGVGVDPFTADYSANVRIORSFDPVVPY  
HFAIVLPDTSPPRVLSEFRALDLALPYETI"  
complement (7308. .8624)  
/gene="tlorf7"  
complement (7308. .8624)  
/gene="tlorf7"  
/function="involved in conjugation"  
/note="Probable trb1 gene"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="BAA87632.1"  
/db\_xref="GI:6498180"  
/translation="MIQSLKGTQAEALSGMRIRNLPILIIIVVLVAFVIGVIE  
GLTSRLYFRDGGPDTSGNPASTPADQIKGVSDGIIIGEPOLVTLQPTPVEIQEE  
KRAVNPFTPOGRREARPEAOLESQVWRARLQREBOYLERBORARLQONDA  
AYDSPIAVDRKFAAGADTAMASATTVGVSATTSASDLVAAALRGAGCONDP  
NGSKKEDFPMADLKEGLYLPNRVYQGSYIELRKGSVYPATLITGINDLPRIHQ  
VSQNVYDSATGRHLIPQGTKLFGYDSVVSFQSKSVLVVWTDIIFPNOSTLDIGGA  
GIDAOQSGSRDKVRKMLQTPSAILLVIGIDMAVPESSLTALDPTASDAARN  
FAETFGVADRTIORMDVQPTLEIRPGKFNVLVDQDIVFPQTVR"  
complement (8637. .9113)  
/gene="tlorf8"  
complement (8637. .9113)  
/gene="tlorf8"  
/function="involved in conjugation"  
/note="Probable trbH gene"

Query Match 89.1%; Score 1605.2; DB 1; Length 206479;  
Best Local Similarity 93.9%; Pred. No. 0; Mismatches 108; Indels 2; Gaps 2;  
Matches 1692; Conservative 0;  
QY 1 ATGTCAAGCTTCACCTCTCCTTGATTAACCAAGTGCATCTCCCAACCAAAATGTGTGAT 60  
Db 188587 ATGTCAAGCTTCACCTCTCCTTGATTAACCAAGTGCATCTCTCTCAAAATGTGTGAT 188646  
QY 61 CTGCAATAGTGCATTAAGCGCGATGAATTGAGCCGACGAGGTTCCGATGCTTCTTAGAA 120  
Db 188647 CTGATATAGTGCATTAAGCGCGATGAATTGAGCCGACGAGGTTCCGATGCTTCTTAGAA 188706  
QY 121 CGAAGACTTCTGAGGAGAGAGATTAATCAAGCTCCACGAGGCGAGCGGTGTTA 180  
Db 188707 CGTGAAGCTTCTGAGGAGAGAGATTAATCAAGCTCCGAGGAGGCGAGCGGTGTTA 188766  
QY 181 GCTTGCAAAAGCGTGCAGTGTGCTCCCTCCCGAGATCTCAGCTGTGAGAAAGTAGCA 240  
Db 188767 GCTTGCAAAAGCGTGCAGTGTGCTCCCTCCCGAGATCTCAGCTGTGAGAAAGTAGCA 188826  
QY 241 GTTCTCTCCGCTTATATCTATATTTGGCAAGAAATTTCTGGGCGGATACTTGAATCGAA 300  
Db 188827 GCCCTCTCCGCTTATATCTATATTTGGCAAGAAATTTCTGGGCGGATACTTGAATCGAA 188886  
QY 301 CCTTGGCGCGGCAACAGAGATGTCTCGTGGCATGAGCTTGAACACCATTTGCATG 360  
Db 188887 CCTTGGCGCGGCAACAGAGATGTCTCGTGGCATGAGCTTGAACACCATTTGCATG 188946  
QY 361 GATTTCTCGAAGACAACTAATCCAAGCCCTTTTGTGTCGAGCGGTAAAGATGTGCA 420  
Db 188947 GATTTCTCGAAGACAACTAATCCAAGCCCTTTTGTGTCGAGCGGTAAAGATGTGCA 189006  
QY 421 CCGATTGATCTTATGATTTTGTGGCCATTTCAATCTTAAGATCTCGGCGCTTTCGAAAC 480  
Db 189007 TCCAGGATCTTATGATTTTGTGGCCATTTCAATCTTAAGATCTCGGCGCTTTCGAAAC 189066  
QY 481 CTGCAATGCGCGTGTACGAGATGCGACGATGAATGCGTTACCGGTTTACCATTAAC 540  
Db 189067 CTGCAATGCGCGCGTGTACGAGATGCGACGATGAATGCGGTTTACCATTAAC 189126

[illegible]

DMGLRRSAKKIDRHNGNAGLGGIPLCFKANIATGIFPTSAATPALINHLPKIPSRV  
AELPLFAGALPGASGMHLESGITSNNTVATGVRWNPMLRIGSSAGGAAVAVSR  
LMIGIGTDTGASVRLPALCGVGRPTIARIPRRIIPVSTRTSTAGIINACVADV  
IILDQVTSRKSATISMPKLRITGLPTTFYDLDLDAFAETTRILIANGVADV  
EADIPIHLEINSGASLPALYEPHALKKYLDVFGVSPDVIKGRSPDVANIVSA  
QIOGHQISNDEVELAROSFRPLQATRYNFRLYQDAILFPAPLAAKAGCESSVI  
HNSMNTFKIYRVNDPSSNAGLPGLSLPACLTIPRPLVGMEDIDGLAGSDHRLAIG  
ALEKXINPSPFPDAPN"  
3151..5418  
/gene="AGR\_pt1\_49"  
3151..5418  
/gene="AGR\_pt1\_49"  
/note="AF126446" tryptophan monooxygenase"  
/codon\_start=1  
/transl\_table=11  
/product="AGR\_pt1\_49p"  
/protein\_id="AAK90969.1"  
/db\_xref="GI:15163473"  
/translation="MSASALLDNOCDFSTKMYDLIMVDKADELDRRVSDFEREAS  
RGRITQISGECSAGLACKRLADGRPEISAGKVAASAIVYVKEILLRIIESEPM  
AKRVSGLVADLAPFCMDSESEKQLOTLTLGSKRCSADLSHPAIVASIKTASRT  
LQMPYEKGTTRKVTGFTLLEAVPFDWVAYGRNMLKASAGSPETIDLYRSF  
DQSDSGRIGFPEPDVKKPVAIIGAGISGLVVASLHAGVDVITYEASDVGGL  
WSHAFKDAPEVVAEMGMRPPAPASCLFFLEERYGSLSPFPNPGTVDNLVYQIGIR  
YMKAGQOPPKLPHRVYSGWRAPLKDFGHDIVLASPVAITOLKSGDIRRAHDSMO  
TWNREGRESSESAIERIFLGTHPGEETSPPHMDLFLMGIISGSGFGPVESEFI  
EITRLVYNGYEBNORMSEGISSELPRIATQVYNGVSGRIRHVOYRAIEKRTKIK  
IRKSGISELYDVVYVTSGLANTQRLRLCTDITIIPAPNOAVDNSHMTGSKLFL  
TERKFWLHDILPSCVLMIDGIAKAVYCLADYPODNGGLVLIYTWEDSDHKLAVD  
KKERFCLIRDAISRSFPAPQALVYPACADYPODNGGLVLIYTWEDSDHKLAVD  
YSEELFQALDMDTNDTGVYLAGCSFTGMEVAGIQTACNAVCAIITHNGGILLAKON  
PLESHMKRYVYRNRN"  
6089..6838  
/gene="AGR\_pt1\_50"  
6089..6838  
/gene="AGR\_pt1\_50"  
/note="Probable tmr gene"  
/codon\_start=1  
/transl\_table=11  
/product="AGR\_pt1\_50p"  
/protein\_id="AAK90970.1"  
/db\_xref="GI:15163474"  
/translation="MYSRYRKLMDLRLIFGPTCTGKTSTAVALAQOTGPLVSLDBV  
OCCPOLSTGSGRPTVEELKGTSLYLDRPLVIGIIIAKQAHRLMEVYVYNAHGL  
ILBGGISILKCAQSSYNSADPRMHIIRHLEADEBETPMVAARVYQMLRPAAGSI  
IOELVDWKSPRLPIKEIDGTRYALPASQNTISDMLQDLADMEDKLINGIAOE  
YLIHARQEQKFRVNAAYDGFGEHFGMY"  
7354..7914  
/gene="AGR\_pt1\_51"  
7354..7914  
/gene="AGR\_pt1\_51"  
/note="Probable 6a gene"  
/codon\_start=1  
/transl\_table=11  
/product="AGR\_pt1\_51p"  
/protein\_id="AAK90971.1"  
/db\_xref="GI:15163475"  
/translation="WTRQGMVLPCLSGKDDQLODELSELSKAYRKRFQTDLTKSD  
IIDPSGFLVYLDEENYRIQROSMVLSNAPDGLATTLLEPVSSTYTFQVBEQQA  
FSGDGRINSRNESSSYFLTIQASNEFGVAVNTFGOSKIDWKRRMPDASOPLD  
YHILAVGSAFLPEALDVEDSDAV"  
8400..8439  
/gene="AGR\_pt1\_bx9p"  
8400..8439  
/gene="AGR\_pt1\_bx9p"  
/note="AGR\_pt1\_bx9p"  
/note="tm1"  
/codon\_start=1  
/transl\_table=11  
/product="AGR\_pt1\_bx9p"  
/protein\_id="AAK90972.1"  
/db\_xref="GI:15163476"  
/translation="MTVANQVRDLTLILRTGEMOSRLQARTEGALLPETYFQPS  
AIRGEFDDDEYILRLQELVYVILREDIARQALRRHLPSNNSNGINAVTAIPLMLMDA

RLNRVMOERCDOGGLVHYQSPHTNOFLAIPNSCFVRFGTUVINNENYGFYARG  
NTBEDEDDDDMDDEDFGAGETRSQGNLINVYIALGSCCHLSA"  
BASE COUNT 2639 a 2177 c 2364 g 2751 t  
ORIGIN  
Query Match 89.0%; Score 1603.6; Db 1; Length 9931;  
Best Local Similarity 93.8%; Pred. No. 0;  
Matches 1691; Conservative 0; Mismatches 109; Indels 2; Gaps 2;  
QY 1 ATGTCAAGCTTCACCTCTCTTGTATACAGTGGATCATCTCCCAACCAAAATGGTGGAT 60  
DB 3151 ATGTCAAGCTTCACCTCTCTTGTATACAGTGGATCATCTCTCAAAATGGTGGAT 3210  
QY 61 CTGACATGTCGATTAAGGCGGATGTAATTGACCCGAGGGTTCCGATGCTCTTGA 120  
DB 3211 CTGATATAGTCGATTAAGGCGGATGTAATTGACCCGAGGGTTCCGATGCTCTCAGAA 3270  
QY 121 CGAAGAGCTTCTAAGGGAAGAGATTACTCAAGCTCAACGAGTGCAGCGCTGGGTTA 180  
DB 3271 CGTGAAGCTTCTAAGGGAAGAGATTACTCAAAATCTCCGGGCGAGTGCAGCGCTGGGTTA 3330  
QY 181 GCTTGCAAAAGCTGCGCGATGTCCTTCCCGAATCTCAGCTGTGAAAGGTAGCA 240  
DB 3331 GCTTGCAAAAGCTGCGCGATGTCCTTCCCGAATCTCAGCTGTGAAAGGTAGCA 3390  
QY 241 GTTCTCGCGCTATATCTATATTGGCAAGAAATCTGGGCGCGATCTTGAATCGAAA 300  
DB 3391 GCCCTTCGCGCTTACATCTATATTGGCAAGAAATCTGGGCGCGATCTTGAATCGAAA 3450  
QY 301 CCTTGGCGCGGCAACAGTACTGTCCTGTCATCGACTTGGACACCATTTGCATG 360  
DB 3451 CCTTGGCGCGGCAACAGTACTGTCCTGTCATCGACTTGGACACCATTTGCATG 3510  
QY 361 GATTTCTCGAAGACACATTAATCCAAAGCCCTTTTCTGCGCGTAAAGATGCA 420  
DB 3511 GATTTCTCGAAGACACATTAATCCAAAGCCCTTTTCTGCGCGTAAAGATGCA 3570  
QY 421 CGGATGATCTTATGTCATTTGCGGCAATTTCAATCTTAAGACTCGCGCTTTGCAACC 480  
DB 3571 TCCAGGATCTTATGTCATTTGCGGCAATTTCAATCTTAAGACTCGCGCTTTGCAACC 3630  
QY 481 CTGCAATGCGCGCTGTACGAAATGCGACGATGAAATGCGTTACCGGATTAACATACC 540  
DB 3631 CTGCAATGCGCGCTGTACGAAATGCGACGATGAAATGCGGATTAACATACC 3690  
QY 541 CTTGAAAGGCGCGCTGCTTTCATGATGCTTAATGCTGAAACCTGATGCTGAAAGGT 600  
DB 3691 CTTGAAAGGCGCGCTGCTTTCATGATGCTTAATGCTGAAACCTGATGCTGAAAGGT 3750  
QY 601 TCGGCAAGGTTCTTTCCAAACAATCGACTCTCAAGACTGACGACCGTTTTCGACCA 660  
DB 3751 TCGGCAAGGTTCTTTCCAAACAATCGACTCTCAAGACTGACGACCGTTTTCGACCA 3810  
QY 661 TGTTCGATATGTAAGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGGAAATGCGG 720  
DB 3811 TGTTCGATATGTAAGGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGGAAATGCGG 3870  
QY 721 GTTATTTGCGCGCTGCTTTCGCAATTCGTCGTGTCGAAACGAATGCTTCAATGCTGGGTA 780  
DB 3871 GTTATTTGCGCGCTGCTTTCGCAATTCGTCGTGTCGAAACGAATGCTTCAATGCTGGGTA 3930  
QY 781 GAGCATGTTACATATATGAAAGCAAGTATCGTGTGGAAGCAAGCTTTGGTCAATGCT 840  
DB 3931 GAGCATGTTACATATATGAAAGCAAGTATCGGCTTGGAGGCAAGCTTTGGTCAATGCT 3990  
QY 841 TTCAGGAGCGCTCTAGTGTCTGTGCGCAATGGGCGCAATGCGATTTCTCTGCTGCA 900  
DB 3991 TTCAGGAGCGCTCTAGTGTCTGTGCGCAATGGGCGCAATGCGATTTCTCTGCTGCA 4050  
QY 901 TTTCTGCTGTTTCTTCTCTGAGCGCTTACGCGCTGCTTCAATGAGCGCTTCCCAAT 960  
DB 4051 TTTCTGCTGTTTCTTCTCTGAGCGCTTACGCGCTGCTTCAATGAGCGCTTCCCAAT 4110

QY 961 CCCGCGACAGTGCACACTTACTGTTGCTTACCAAGCGCTGCATATCATGTGAAAACCGGG 1020  
 DB 4111 CCCGCGACAGTGCACACTTACTGTTGCTTACCAAGCGCTGCATATCATGTGAAAACCGGG 4170  
 QY 1021 CACTGCGACCGAAGCTGTTCCATCGGTTTACAAGCGTTGGCGGTTCTTGAAGGAC 1080  
 DB 4171 CACAGCGACCGAAGCTGTTCCATCGGTTTACAAGCGTTGGCGGTTCTTGAAGGAC 4230  
 QY 1081 GATTTTCATAGAGAGATATGTTGTTGGCTTGGCTTGGCTTACTTACTGAGCGCTTGA 1140  
 DB 4231 GATTTTCATAGAGAGATATGTTGTTGGCTTGGCTTGGCTTACTTACTGAGCGCTTGA 4290  
 QY 1141 TCAGACACATTAAGTGGGCTCATGATCTCTGGCAATTTGGCTGAACCGTTTCGGAGG 1200  
 DB 4291 TCAGACACATTAAGTGGGCTCATGATCTCTGGCAATTTGGCTGAACCGTTTCGGAGG 4350  
 QY 1201 GATTCCTTCTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGTTGGT-AA 1259  
 DB 4351 GATTCCTTCTTCAGGAGATAGAGAGATCTTTCTGGGCAACATCTCTGTTGGTAA 4410  
 QY 1260 ACATGAGTTTCTCTCATGATGAGGACCTATTCAGCTATAGGGAATAGGATCTGGCGG 1319  
 DB 4411 ACATGAGTTTCTCTCATGATGAGGACCTATTCAGCTATAGGGAATAGGATCTGGCGG 4469  
 QY 1320 GTTTCGTCAGTTTTCGAAAGCGGTTTATTTAGATCTCTCGCTGCTCATCAACGGATA 1379  
 DB 4470 GTTTCGTCAGTTTTCGAAAGCGGTTTATTTAGATCTCTCGCTGCTCATCAACGGATA 4529  
 QY 1380 TGAAGAAATCAGCGGATGTCCTCGTGAAGAAATCTCAGAACTTCACGTCGATTCGCATC 1439  
 DB 4530 TGAAGAAATCAGCGGATGTCCTCGTGAAGAAATCTCAGAACTTCACGTCGATTCGCATC 4589  
 QY 1440 TGAAGTTTAAAGGTTGCTGTTGAGCGCGCATATGCGCATGTTCAATCGCGCGAT 1499  
 DB 4590 TGAAGTTTAAAGGTTGCTGTTGAGCGCGCATATGCGCATGTTCAATCGCGCGAT 4649  
 QY 1500 TCAGAGAAAGAAACAAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTATGA 1559  
 DB 4650 TCAGAGAAAGAAACAAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTATGA 4709  
 QY 1560 TAAAGTGTGTGTCATCTGAGCTCGCAAAATTCACATCGAGCATTCCTGATCGCA 1619  
 DB 4710 TAAAGTGTGTGTCATCTGAGCTCGCAAAATTCACATCGAGCATTCCTGATCGCA 4769  
 QY 1620 TACCAATATTTTACGCGACCACTGAAACGAGCGTTGAATACACCATATGACAGATC 1679  
 DB 4770 TACCAATATTTTACGCGACCACTGAAACGAGCGTTGAATACACCATATGACAGATC 4829  
 QY 1680 GTCAAACTTCTCTGATGATGACGAAATTTCTGGTTAGACCATATCTCCCGTCTTG 1739  
 DB 4830 GTCAAACTTCTCTGATGATGACGAAATTTCTGGTTAGACCATATCTCCCGTCTTG 4889  
 QY 1740 TGTCTCATGAGCGGATTCGCAAAAGCATGTAATGCTTGAGTATGAGTTCGAGATCC 1799  
 DB 4890 TGTCTCATGAGCGGATTCGCAAAAGCATGTAATGCTTGAGTATGAGTTCGAGATCC 4949  
 QY 1800 GA 1801  
 DB 4950 GA 4951

RESULT 9  
 AEO09419 10377 bp DNA linear BCT 20-DEC-2001  
 LOCUS Agrobacterium tumefaciens str. C58 Ti plasmid, section 2 of 21 of  
 DEFINITION the complete sequence.  
 ACCESSION AEO09419 AEO08690  
 VERSION AEO09419.1 GI:17744019  
 KEYWORDS Agrobacterium tumefaciens str. C58 (U. Washington)  
 ORGANISM Agrobacterium tumefaciens str. C58 (U. Washington)  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

## REFERENCE

1 (bases 1 to 10377)  
 AUTHORS Wood,D.W., Seubal,J.C., Kaul,R., Monks,D., Chen,L., Wood,G.E.,  
 Zhou,Y., Woo,L., Kitajima,J.P., Okura,V.K., Almeida Jr.,N.F.,  
 Zhou,Y., Boye Sr.,D., Chapman,P., Clendenning,J., Deatherage,G.,  
 Gillet,M., Grant,C., Guenther,D., Kueyavin,T., Levy,R., Li,M.,  
 McClelland,E., Palmeri,A., Raymond,C., Rouse,G.,  
 Saenphimmachak,C., Wu,Z., Gordon,D., Eisen,J.A., Paulsen,I.,  
 Karp,P., Romero,P., Zhang,S., Yoo,H., Tao,Y., Biddle,P., Jung,M.,  
 Krespan,M., Perry,M., Gordon-Kamm,B., Liao,L., Kim,S., Hendrick,C.,  
 Zhao,Z., Dolan,M., Tingey,S.V., Tomb,J., Gordon,M.P., Olson,M.V.,  
 and Nester,E.W.  
 TITLE The genome of the natural genetic engineer Agrobacterium  
 tumefaciens C58  
 JOURNAL Science 294 (5550), 2317-2323 (2001)  
 MEDLINE 21608550  
 PUBMED 11743193  
 REFERENCE 2 (bases 1 to 10377)  
 Wood,D.W., Seubal,J.C., Kaul,R., Monks,D., Chen,L., Wood,G.E.,  
 Chen,Y., Woo,L., Kitajima,J.P., Okura,V.K., Almeida Jr.,N.F.,  
 Zhou,Y., Boye Sr.,D., Chapman,P., Clendenning,J., Deatherage,G.,  
 Gillet,M., Grant,C., Guenther,D., Kueyavin,T., Levy,R., Li,M.,  
 McClelland,E., Palmeri,A., Raymond,C., Rouse,G.,  
 Saenphimmachak,C., Wu,Z., Gordon,D., Eisen,J.A., Paulsen,I.,  
 Karp,P., Romero,P., Zhang,S., Yoo,H., Tao,Y., Biddle,P., Jung,M.,  
 Krespan,M., Perry,M., Gordon-Kamm,B., Liao,L., Kim,S., Hendrick,C.,  
 Zhao,Z., Dolan,M., Tingey,S.V., Tomb,J., Gordon,M.P., Olson,M.V.,  
 and Nester,E.W.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-SEP-2001) Department of Microbiology, University of  
 Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA  
 98195-7242, USA

## FEATURES

source 1.10377  
 Location/Qualifiers  
 /organism="Agrobacterium tumefaciens str. C58 (U.  
 Washington)"  
 /mol\_type="genomic DNA"  
 /strain="C58"  
 /db\_xref="taxon:180835"  
 /plasmid="Ti"  
 319..513  
 /gene="ipt"  
 /note="synonym: Atu6006"  
 319..513  
 /gene="ipt"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer"  
 /codon\_start=1  
 /transl\_table=1  
 /product="isopenityl adenosine phosphotransferase"  
 /protein\_id="AAL46242.1"  
 /db\_xref="GI:17744020"  
 /translation="MDLVLFEGPTGKTAATLSEHTGIPVLAIDRIQCPQSLSTG  
 SGRPAAEYTAPOSIGQKRC"  
 complement(380..1414)  
 /gene="mas1"  
 /note="synonym: Atu6007"  
 complement(380..1414)  
 /gene="mas1"  
 /note="identified by sequence similarity; putative; ORF  
 located using Blastx/Glimmer"  
 /codon\_start=1  
 /transl\_table=1  
 /product="agrocinnopine synthase reductase"  
 /protein\_id="AAL46243.1"  
 /db\_xref="GI:17744021"  
 /translation="MHEGLGPRKMFENVPPCEKTELFSARVAGLTHARRNILEV  
 SHGVLVAVATLIGVETAMORCNACVLFKFDGSMVYHTDSVILVTGASGGIDK  
 AIDNLMRGRGRLSGARSVEKTEATPRQNESLFPARFEADLTQMEWVSTAMIKF  
 GRIDGLVNNAGDTRVALAEGIDYDLEROMKINAGAPURKIMLCLPQULAKSGSGRIV  
 NIISSGQRLVNSFVGVMWTKALSLGLTKTAQHVGESEGVRTMEICPGVATKMSAMT  
 DEIEDDGIOEPHIELVROAIERDNNAFVPEKIEVLICIQPPLAVRSRSKVSIVYGP  
 KRPCCQCAH"  
 complement(1622..2878)

gene

gene  
/gene="acex"  
/note="synonym: Atu6008"  
complement (1622. .2878)  
/gene="acex"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=1  
/product="agrocinnopine synthase"  
/protein\_id="AAL46244.1"  
/db\_xref="GI:17744022"  
/translation="MEGTVPAPWTLMWDLPEVNTPEVLARSHSPVLLERLEEDLPLO  
VIRHGWFLNGKRIQECTASLLAALQGGRLNSELVCLTSDNVAIVSHDLNTWVS  
EKLGDKLFNEHSSKIDVPIREVSNGIIOKRYLETIDHILPTEIFSKVLANS  
ATIFLDGRNEAVHIVAMLSHREPHORVVVLFYEPYHGAFLVDAVLAQASAMR  
KSIAMALPRELCLARROYTEPVDLYLAKGMTMRAKRPHLKRAVYRCIDFPAAL  
VTNLLGQVVDKVLAFDSQNAVRLAYTLKEDTMRARPHLKRAVYRCIDFPAAL  
LDSGERGEFSIDIKTGRARHETDERKHIRKRGTPGNSATYADWVLSDRPEDEMAIW  
EMRNOGIDREVSHSLDLNLETSK"  
3636. .4157  
/gene="genes5"  
/note="synonym: Atu6009"  
3636. .4157  
/gene="genes5"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=1  
/product="gene 5 protein"  
/protein\_id="AAL46245.1"  
/db\_xref="GI:17744023"  
/translation="MNSIINTDVPIDPAIDEVVRCFEVACLPAGAPILNLLNSL  
TVYVCSQAMRKXAKHAFYDVGSDGVISTVPAVAGITKEHMRSHNNVCONTSSE  
THOLDATIAFLPTSLQNPSPSHMKIGCDSFLASRNDPCEVILAAVKGALFHNGLKK  
NPRCGGWLPPSC"  
complement (4662. .6065)  
/gene="tms2"  
/note="synonym: Atu6010"  
complement (4662. .6065)  
/gene="tms2"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=1  
/product="indole acetamide hydrolase"  
/protein\_id="AAL46246.1"  
/db\_xref="GI:17744024"  
/translation="WVPITSLAQTLELRKRDVSCLEVETLIARCOAKPLNALLAT  
DMGLRSAKRIDRHGNAGLGLGIPLCFANATGIFPTSAATPALINHLPIPSRV  
AERLFSAGALPGASGNHLSFGITSNVATGAVRNPMBSLIPGSSGGVAAVASR  
LMIGGIGTDGASVRLPALCGVGPPTLARYPRDIIIPSPTRDTAGIIAACVADV  
IILDVYISGRSAKISMPKLGRLGRLPTTFYDDLDADVAPAAETTRILANGVPTV  
EADIPHLBELNLSALPILAYEPHALKXLDVGVSVSDVIKGRSPDVANIVSA  
QIOGHQISNDEVELARQSPRPLQATYRNFRLYQDALIFPTAPLAAKAIGESSVT  
HNSGMNTFKIYRNVDPSSNAGLPGSLPACLPRLPVGMEIDGLAGSDHRLAIG  
AALEKAINPSPDAFN"  
6403. .8670  
/gene="tms1"  
/note="synonym: Atu6011"  
6403. .8670  
/gene="tms1"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=1  
/product="tryptophan 2-monooxygenase"  
/protein\_id="AAL46247.1"  
/db\_xref="GI:17744025"  
/translation="MSASALIDNOCDFSTKMDVLIWMDKADELDREVSAPRSERAS  
RGRRTIISGECSAGLACKRLADRPPEIISAGEVYALSIYIGKILIRLIESBEM  
ARRAVSGVLAIDLAPCMDFSEADQLQTLFLSGKRASSDLSHFVAISISKARST  
LQMPYEKGTTRKVTGTLTLEAVPDPMAVGNLMLKASGSPPTIDLLYRSFP  
DQSDSGRIGFFPEDVKKPAIIGAGISGLVAVSELHAGVSDVTIYEASDRVGKL

MSHAFKDAFVVAEMGAMRPPAASCLFFELERYGLSSMRPFNPCTVDNLVYQGLR  
YMKAGQCPKLFHRYVSGMARFLKDGFEHGIIVLASPAVIAQALKSGDIRAHDSWO  
TWNPRGREGFSFAIERILPGLTHPPRDEHIDFLKMGISGGGPGVPEESGFI  
EILRLVINGEENORMCSHGISLPRRIATOVNGVSVORIRHVORAIIEKTKIK  
ILKSGISLRYKRVVTTSLANQLRHLCTCTOTTIFRAVNAVNDMSHTSSKIFLL  
TERKEWLDHILPSCVMDGIARVCLDIEPDGPKGLVLSYEDDSHKLAVPD  
KKEKFCFLDAISRSPAPAOHLVPAQCAVDNVQHDMLTBNKGAFLKMRGSDP  
VSEELFFQALDMTNDTGVYLAAGSCSFTGWVEGAIQTACNAVCAIINHCGILAKDN  
PLEHSWKRYNRYRN"  
9368. .10090  
/gene="lpc"  
/note="synonym: Atu6012"  
9368. .10090  
/gene="lpc"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=1  
/product="isopentenyl transferase"  
/protein\_id="AAL46248.1"  
/db\_xref="GI:17744026"  
/translation="MDLRILGPTGKTSTAVALAONGLPVLSLDRVOCPSLSTG  
SGRPYBELKGTSLYLDPRPLVKGIIIAKQHERLMGVSYYEARAGLIEGGSISL  
LCKMAQSSYWSADFRWHIIRHLEADETFMNAKARVQMLPAGLSTIIQELVDLAK  
EHLRLILKEIDGYRYAMPASQONITSDMLQLDADMDKLIHGAOBYLLIARROE  
OKFRVNAAYADGFEHGFPMGY"  
BASE COUNT 2745 a 2376 c 2438 g 2818 t  
ORIGIN

Query Match 89.0%; Score 1603.6; DB 1; Length 10377;  
Best Local Similarity 93.8%; Pred. No. 0;  
Matches 1691; Conservative 0; Mismatches 109; Indels 2; Gaps 2;

QY 1 ATGTGAGCTTCACTCTCTGATTAACGAGTGATCATCTCCCAACCAAAATGATGAT 60  
|||||  
DB 6403 ATGTGAGCTTCACTCTCTGATTAACGAGTGATCATCTCTCAACCAAAATGATGAT 6462  
61 CTGACATAGTGCATTAAGCGCGATGATGATGACCGGAGGATTCGATGCTTTAGAA 120  
|||||  
DB 6463 CTGATTAATGATGATTAAGCGCGATGATGATGACCGGAGGATTCGATGCTTTAGAA 6522  
121 CGAAGACTTTTAAAGGAAAGAGATTTACTCAAGCTCCACGAGTGCAGCGCTGGTTA 180  
|||||  
DB 6523 CTGGAAGCTTCTTGGGGAAGAGATTTACTCAAGCTCCGCGAGTGCAGCGCTGGTTA 6582  
181 GCTTGAAGAAAGCTGCGCATGCTGCTTCCCGAGATCTCAGCTGTGGAAGATGACA 240  
|||||  
DB 6583 GCTTGAAGAAAGCTGCGCATGCTGCTTCCCGAGATCTCAGCTGTGGAAGATGACA 6642  
241 GTTCTCTCGCTTATATCTATATTTGGCAAGAAATCTGGGCGGATTAATTGAGAA 300  
|||||  
DB 6643 GCCCTCTCCGCTTACATCTATATTTGGCAAGAAATCTGGGCGGATTAATTGAGAA 6702  
301 CTTTGGCGCGGCAACAGTGAAGTCTCTGTTGCCATGATGATGACACCATTTGCATG 360  
|||||  
DB 6703 CTTTGGCGCGGCAACAGTGAAGTCTCTGTTGCCATGATGATGACACCATTTGCATG 6762  
361 GATTTCTCGAAGCAACATCTTCCAAACCTGTTTTGTGAGCGGTAAAGATGACA 420  
|||||  
DB 6763 GATTTCTCGAAGCAACATCTTCCAAACCTGTTTTGTGAGCGGTAAAGATGACA 6822  
421 CCGATTGATTTAGTATTTCTGAGCAATTTCAATCTTAAGATCGCGCTTTTCAACC 480  
|||||  
DB 6823 TCCAGGATTTAGTATTTCTGAGCAATTTCAATCTTAAGATCGCGCTTTTCAACC 6882  
481 CTGCAATGCGCGGTGACGAAATGCGACGATGAATGCGGTTAACCATTAACC 540  
|||||  
DB 6883 CTGCAATGCGCGGTGACGAAATGCGACGATGAATGCGGTTAACCATTAACC 6942  
541 CTTGAAGGCGCGGTGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
|||||  
DB 6943 CTTGAAGGCGCGGTGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 7002

QY 601 TCGCAGAGTTCCTTCCAAATCGACTGCTCTACGACTGACAGCCGTTTTTGAACAA 660  
Db 7003 TCGGCAAGTTCCTTCCAAATCGACTGCTCTACGACTGACAGCCGTTTTTGAACAA 7062  
QY 661 TGTTCGATAGTGAAGGATCGGCTTTTCCGAGAGATGTTCTTAAGCGAAATGGCG 720  
Db 7063 TGTTCGATAGTGAAGGATCGGCTTTTCCGAGAGATGTTCTTAAGCGAAATGGCG 7122  
QY 721 GTGATGGGCGCTGGCATTTCCGAGCTGCTGGTGGCGAAACGATGCTTCAATGCTGGGTA 780  
Db 7123 ATGATGGGCGCTGGCATTTCCGAGCTGCTGGTGGCGAAACGATGCTTCAATGCTGGGTA 7182  
QY 781 GACGATGTTCAATATATGAAGCAAGTATGCTGTGAGGCAAGCTTGTGATCATGCT 840  
Db 7183 GAGATGTTCAATATATGAAGCAAGTATGCTGTGAGGCAAGCTTGTGATCATGCT 7242  
QY 841 TTGAGGAGCGCTCTAGTGTGTCGGCCGAAATGGGGCGATGCGATTTCTCTGCTGCA 900  
Db 7243 TTGAGGAGCGCTCTAGTGTGTCGGCCGAAATGGGGCGATGCGATTTCTCTGCTGCA 7302  
QY 901 TTCTGCTGTTTTTCTCTCTCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
Db 7303 TCTGCTGTTTTTCTCTCTCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7362  
QY 961 CCGGCAAGTGCACACTTACTTGTCTACCAAGCGTCCAAATCATGTGAAAGCGGG 1020  
Db 7363 CCGGCAAGTGCACACTTACTTGTCTACCAAGCGTCCAAATCATGTGAAAGCGGG 7422  
QY 1021 CAGCTGCAAGCTGTTTCAATCGGTTTACCAAGCGGTTGGCGGTTCTTGAAGAC 1080  
Db 7423 CAGCTGCAAGCTGTTTCAATCGGTTTACCAAGCGGTTGGCGGTTCTTGAAGAC 7482  
QY 1081 GGTTCATGAGGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 7483 GGTTCATGAGGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7542  
QY 1141 TCAAGCACATTAAGTGGGCTCATGACTCTGGCAATTTGGCTGAAACGTTTCGGAGG 1200  
Db 7543 TCAAGCACATTAAGTGGGCTCATGACTCTGGCAATTTGGCTGAAACGTTTCGGAGG 7602  
QY 1201 GAGTCTCTCTCTTCAAGGATAGAGAGATCTTCTGGGCAACATCTCTGCTGCTGCT 1259  
Db 7603 GAGTCTCTCTCTTCAAGGATAGAGAGATCTTCTGGGCAACATCTCTGCTGCTGCTGCT 7662  
QY 1260 ACATGAGTCTCTCTCATGATGAGGACTTCAAGCTTATGGAATGAGATCTGGCGG 1319  
Db 7663 ACATGAGTCTCTCTCATGATGAGGACTTCAAGCTTATGGAATGAGATCTGGCGG 7721  
QY 1320 GTTGTGCTAGTTTGAAGCGGGTTATGAGATCTCTGGCTGCTGCTGCTGCTGCTGCT 1379  
Db 7722 GTTGTGCTAGTTTGAAGCGGGTTATGAGATCTCTGGCTGCTGCTGCTGCTGCTGCTGCT 7781  
QY 1380 TGAAGAAATCAGCGATGTCCTGAGGAAATCTCAGAACTTCAAGCTGATCGCATC 1439  
Db 7782 TGAAGAAATCAGCGATGTCCTGAGGAAATCTCAGAACTTCAAGCTGATCGCATC 7841  
QY 1440 TGAAGTGTAAAGGCTGCTGCTGAGCCGCGCATATGCCATGTTCAATCAGGCGAT 1499  
Db 7842 TGAAGTGTAAAGGCTGCTGCTGAGCCGCGCATATGCCATGTTCAATCAGGCGAT 7901  
QY 1500 TCAAGGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGGGGATATCTGAACTTATGA 1559  
Db 7902 TCAAGGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGGGGATATCTGAACTTATGA 7961  
QY 1560 TAAAGTGTGTGATCAATCTGAGCTCGCAATATCAATCAGGCTGCTGATCGCA 1619  
Db 7962 TAAAGTGTGTGATCAATCTGAGCTCGCAATATCAATCAGGCTGCTGATCGCA 8021  
QY 1620 TACCAATATTTTCAAGCAGCAAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 1679  
Db 8022 TACCAATATTTTCAAGCAGCAAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 8081  
QY 1680 GTCAAAACCTTCTCTGATGATGACGAAATTTCTGCTTGAACATATCTCTCCGCTTTC 1739

Db 8082 GTCAAAACCTTCTCTGATGATGACGAAATTTTGTGATGACCATATCTCTCCGCTTCG 8141  
QY 1740 TGTTCCTCATGAGCGGATCGCAAAAGCAAGTATGCTCTGAGTATGATCGCAGATCC 1799  
Db 8142 TGTTCCTCATGAGCGGATCGCAAAAGCAAGTATGCTCTGAGTATGATCGCAGATCC 8201  
QY 1800 GA 1801  
Db 8202 GA 8203  
RESULT 10  
AF126446  
LOCUS  
DEFINITION  
AF126446 5968 bp DNA linear BCT 25-FEB-2000  
Agrobacterium tumefaciens plasmid pTiC58 agrocinopine synthase  
(acs) gene, partial cds; and 5 protein (5'), indole-3-acetamide  
hydrolase (iaah), and tryptophan monooxygenase (iaam) genes,  
complete cds.  
ACCESSION  
AF126446  
VERSION  
AF126446.1 GI:4836565  
KEYWORDS  
SOURCE  
ORGANISM  
Agrobacterium tumefaciens (Rhizobium radiobacter)  
Agrobacterium tumefaciens  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
REFERENCE  
1 (bases 1 to 5968)  
Olsen, L., Salomone, J. Y., Helfer, A., Schmidt, J., Hammann, P. and De  
Ruffray, P.  
Sequence and functional analysis of the left-hand part of the  
T-region from the nopaline-type Ti plasmid, pTiC58  
Plant Mol. Biol. 41 (6), 765-776 (1999)  
JOURNAL  
MEDLINE  
PUBMED  
10737141  
2 (bases 1 to 5968)  
Olsen, L., De Ruffray, P. and Hammann, P.  
Direct Subregion  
Submitted (08-FEB-1999) Cell Biology, IBMP, rue du General Zimmer  
12, Strasbourg 67804, France  
LOCATION/Qualifiers  
FEATURES  
source  
1..5968  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/strain="C58"  
/db\_xref="taxon:358"  
/plasmid="pTiC58"  
/note="Ti plasmid; T-DNA"  
complement(1..19)  
/gene="acs"  
complement(1..19)  
/gene="acs"  
/codon\_start=1  
/transl\_table=1  
/product="agrocinopine synthase"  
/protein\_id="AAD30486.1"  
/db\_xref="GI:4836566"  
translation="MFNIGK"  
782..1462  
/gene="5"  
782..1462  
/gene="5"  
/function="may catalyze synthesis of indole-3-lactate"  
/codon\_start=1  
/transl\_table=1  
/product="5 protein"  
/protein\_id="AAD30487.1"  
/db\_xref="GI:4836567"  
translation="MYHSRPPIIIDSNIQDRBLKVLRTETAYRSPADLIPA  
QKSWNSITNDVPIIDPAIDVYKRFCEVACLPAGIPLNITLNDLSLYVCSFAM  
RKYAKRFYDGVSDGVISTVPAEGLTKETMSMHNVCQNTSNETHDIDATYAF  
LPTSLQNSFSHMKIGCDSFLAPSRVDPCVEIIVGKALFHDNGLKKPKWMSAP  
TFLKRLV"  
complement(1953..3353)  
gene

CDS  
/gene="iaaH"  
complement (1953 . .3353)  
/gene="iaaH"  
/function="synthesizes indole-3-acetic acid from  
indole-3-acetamide"  
/codon\_start=1  
/transl\_table=11  
/product="indole-3-acetamide hydrolase"  
/protein\_id="A03048.1"  
/db\_xref="GI:4836568"  
/translation="MPTISLAQTLERKRYSCLELVKTLARCOAEPNALAT  
DMGLRRTAKIDRHGAGLGLGIPLCFEGNATATGIPSAATPALINLPISRV  
POTLPSAGLPAGSNMHELSFGITSNVATGAVRNPWPSLPGSSGAPAAVSR  
LMIGICTDGAASVRLPALCGVGFPTLARYPRDIIPVSPRTDAGIIACVADV  
IILDVIGSRGSAKILPMLKGRIGLPTTYEDDLDAVDVAFAETIRLARGVFEV  
EADIPLHEELNSGASLPIALYEPHALIKYLDPEVGFELFETORISRPVANVASO  
IDHJOISNDEYELAROSFRPLDATTNTRFRLYOLDIIFPTPLAKAIGOSSYTH  
NGSMWMTFKIYVRNVDPSNAGLPGLSPLCLTPRLPLVGMEDLAGSDHRLAIGA  
ALBKAINFSPSPAFN"  
3690 . .5957  
/gene="iaaM"  
3690 . .5957  
/gene="iaaM"  
/function="synthesizes indole-3-acetamide from tryptophan"  
/codon\_start=1  
/transl\_table=11  
/product="tryptophan monooxygenase"  
/protein\_id="A03048.1"  
/db\_xref="GI:4836569"  
/translation="MSASALDNDQCHDFSTKWDLIMVDADELDRRVSDAFSERENS  
RGRMTQISEGCSAGLAKRLADGRPELISAGEKVAASVYIVGKEILRIEISEW  
ARAVSGLVAILDALPCMDPSEADLQTLTLGSKRASSDLHSKTSRST  
LQMPRYKGTTRKVTGFTLLEAVPDMVAYGRNMLKASGSPFIDLYRSPF  
DOCSDSGRIGFEPEDVPKVALIAGISGLVVASLHAGVDPTVYASDVVGL  
WSAFLDABSVVAMKAMRPPRPAASCLFFLEKYGSSNRPFPNPTVDNLYQGR  
TYMRAGQPPKLFHRVYSGWRAPLKDGFHGDVLASPAVITALSGDIRAHDSQ  
TINRGRSFPSSAIERIPLGTHPFGGETSPHMDLFLMIGISGSGFGEVSEGI  
EILRLVNGEYENQRMCSGISELPRIATQVNVGSORIRHVVQALAEKRTFK  
IRKSGISELYDKVYVTSGLANTOLRHCLDITIPAPVNOAVDASHMGSKFLK  
TERKFMILDLPSCLVMDGIAKAVYCLDVPDOPNGKGYLISYTMEDSDHKLAVD  
KERFCLDLDAISRSFPAPFAOHLVPACADVDQNVVQDITDENAGAFULNRGDF  
YSEELFQALDMTNDGVYLAGSCSFTGWEVGAIQTACNAVCAIINHNGGILADON  
PLEHSMKRYVYRNRN"  
BASE COUNT 1543 a 1333 c 1472 g 1620 t  
ORIGIN  
Query Match 89.0%; Score 1602; DB 1; Length 5968;  
Best Local Similarity 93.8%; Pred. No. 0;  
Matches 1690; Conservative 0; Mismatches 110; Indels 2; Gaps 2;  
QY 1 ATGTCAAGCTTCACTCTCTTGATTAACGAGGCGATCATCTCCCAACCAAAATGGTGAT 60  
DB 3690 ATGTCAAGCTTCACTCTCTTGATTAACGAGGCGATCATCTCTCAACCAAAATGGTGAT 3749  
QY 61 CTGACAAATGATCGATTAAGCGGATGAATTGACCGGAGGTTTCCGATGCTTTAGAA 120  
DB 3750 CTGATAATGATCGATTAAGCGGATGAATTGACCGGAGGTTTCCGATGCTTTAGAA 3809  
QY 121 CGAGAACTTCTAGGGAAGAAGATTACTCAAAAGCTTCCACCGAGTGCAGCGCTGGGTTA 180  
DB 121 CGAGAACTTCTAGGGAAGAAGATTACTCAAAAGCTTCCACCGAGTGCAGCGCTGGGTTA 3869  
QY 181 GCTTGAAGAGGCGGCGATGATGCTTCCCGAGATCTCAGTGTGTGAAAGGTGCA 240  
DB 3870 GCTTGAAGAGGCGGCGGCGATGATGCTTCCCGAGATCTCAGTGTGTGAAAGGTGCA 3929  
QY 241 GTTCTCTCCGCTTATCTATTTGCAAGAAATTTGSGGCGGATCTGTAATCGAAA 300  
DB 3930 GCTTCTCTCCGCTTATCTATTTGCAAGAAATTTGSGGCGGATCTGTAATCGAAA 3989  
QY 301 CTTTGGGCGGCGGACAGTGAAGTGTCTCTGTTGCCATCGACTTGAACACCATTTGCATG 360  
DB 3990 CTTTGGGCGGCGGACAGTGAAGTGTCTCTGTTGCCATCGACTTGAACACCATTTGCATG 4049

QY 361 GATTTCTCGAAGCACAACTAATCCAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGA 420  
DB 4050 GATTTCTCGAAGCACAACTTCTCCAAACCTGTTTTGCTGAGCGGTAAAGATGTGA 4109  
QY 421 CCGATTGATCTTAATGTCATTTCTGTCGCTTCAATCTTAAGCTGCGGCTTTCCAAAC 480  
DB 4110 TCCAGGATCTTAATGTCATTTCTGTCGCTTCAATCTTAAGCTGCGGCTTTCCAAAC 4169  
QY 481 CTGCGAATGCGGCTTAAGAGAAATGCGAGTAAATGGTTTCCGGGTTTACATTAAC 540  
DB 4170 CTGCGAATGCGGCTTAAGAGAAATGCGAGTAAATGGTTTCCGGGTTTACATTAAC 4229  
QY 541 CTTGAAGGCGGCTGCGATTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB 4230 CTTGAAGGCGGCTGCGATTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 4289  
QY 601 TCGGAGGTTCTTTCCAAACATCGACTGCTCTACGACTGACAGCCGTTTTTGACAA 660  
DB 4290 TCGGAGGTTCTTTCCAAACATCGACTGCTCTACGACTGACAGCCGTTTTTGACAA 4349  
QY 661 TGTTCCGATAGTGAAGGATCGGCTCTTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 720  
DB 4350 TGTTCCGATAGTGAAGGATCGGCTCTTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 4409  
QY 721 GTCAATGCGGCTGCGATTTCCGAGCTCGTGTGCGAAACGAACTGCTTCAATGCTGGGATA 780  
DB 4410 GTCAATGCGGCTGCGATTTCCGAGCTCGTGTGCGAAACGAACTGCTTCAATGCTGGGATA 4469  
QY 781 GACGATGTTCAATATATGAGAGCAAGTATCGTGTGAGAGCAAGCTTTGTCATACGT 840  
DB 4470 GACGATGTTCAATATATGAGAGCAAGTATCGTGTGAGAGCAAGCTTTGTCATACGT 4529  
QY 841 TTGAGGAGGCTCTGATGTCGAGGCGAAATGGGGCGCATGCGATTTCTCTGCTGCA 900  
DB 4530 TTGAGGAGGCTCTGATGTCGAGGCGAAATGGGGCGCATGCGATTTCTCTGCTGCA 4589  
QY 901 TTTGCTGTTTTTTCTTCTCGAGCTTACGAGCTGCTTCTGATGAGGCGGTTCCAAAT 960  
DB 4590 TTTGCTGTTTTTTCTTCTCGAGCTTACGAGCTGCTTCTGATGAGGCGGTTCCAAAT 4649  
QY 961 CCGGCAAGTGCACCTTAATTGTTGCTTCAAGAGCGTCCAAATATGTTGAAAGCCGGG 1020  
DB 4650 CCGGCAAGTGCACCTTAATTGTTGCTTCAAGAGCGTCCAAATATGTTGAAAGCCGGG 4709  
QY 1021 CAGCTGCAACCGAGCTTCCATCGCGTTTCAAGAGTGGGCTGCTTGAAGGAC 1080  
DB 4710 CAGCTGCAACCGAGCTTCCATCGCGTTTCAAGAGTGGGCTGCTTGAAGGAC 4769  
QY 1081 GATTTTCATGAGCGAGATATTTGTTGCTTCCGCTGCTGCTTACTCAGGCTTTGAAA 1140  
DB 4770 GATTTTCATGAGCGAGATATTTGTTGCTTCCGCTGCTGCTTACTCAGGCTTTGAAA 4829  
QY 1141 TCAGAACATTAAGTGGGCTCATGACTCTGCGAAATTTGCTGAAACGTTTGGGAGG 1200  
DB 4830 TCAGAACATTAAGTGGGCTCATGACTCTGCGAAATTTGCTGAAACGTTTGGGAGG 4889  
QY 1201 GAGTCCCTTCTTCAAGGATAGAGAGATCTTTCTGAGGACACATCTCTGAGTGGT-AA 1259  
DB 4890 GAGTCCCTTCTTCAAGGATAGAGAGATCTTTCTGAGGACACATCTCTGAGTGGT-AA 4949  
QY 1260 ACATGAGATTTTCTCATGATTTGGAACCTTAATCAACTAATGGAATAGAGTGGCGGG 1319  
DB 4950 ACATGAGATTTTCTCATGATTTGGAACCTTAATCAACTAATGGAATAGAGTGGCGGG 5008  
QY 1320 GTTGTGCAAGTTTGAAGCGGTTTATGAGATCTCTCGTTGCTCAACCGATA 1379  
DB 5009 GTTGTGCAAGTTTGAAGCGGTTTATGAGATCTCTCGTTGCTCAACCGATA 5068  
QY 1380 TGAAGAAATCAAGCGGATGTCCTTGAAGAACTCAGAACTTCCAGCTGCGATCGCATC 1439  
DB 5069 TGAAGAAATCAAGCGGATGTCCTTGAAGAACTCAGAACTTCCAGCTGCGATCGCATC 5128





QY 841 TTCAGGAGCCTCTAGTGTGTGTCGCGAAATGGGGCGATGCGATTTCTCTGTCTGCA 900  
DB 20827 TTCAAGGACGCTCCACAGCTGTGTGCGGAATGGGGCGATGCGATTTCTCTGTCTGCA 20886  
QY 901 TTCTGCTTTTCTTCTCTCTGAGCGTTAGCGCTGTCTTCTGATGAGCGCGTTCCCAAT 960  
DB 20887 TCGGCTGTGTTTCTCTCTGAGCGGTAGCGCTGTCTTCTGATGAGCGCGTTCCCAAT 20946  
QY 961 CCGGCGACAGTCGACCTTACTTGGTGTACCAAGCGCTCCCAATCATGTGGAAGCGCGG 1020  
DB 20947 CCGGCGACAGTCGACCTTACTTGGTGTACCAAGCGCTCCCAATCATGTGGAAGCGCGG 21006  
QY 1021 CAGCTGCCACCGAAGCTTTCATCGCGTTTCAACACCGTTGGCGTGGCTTTGAAGAC 1080  
DB 21007 CAGAGCGACCGAAGCTTTCATCGCGTTTCAACACCGTTGGCGTGGCTTTGAAGAC 21066  
QY 1081 GGTTCATAGCGAGATATTGTGTGGCTTCCGCTGTCCCTATTTACTCAGCGCTTGA 1140  
DB 21067 GGTTCATAGCGAGATATTGTGTGGCTTCCGCTGTCCCTATTTACTCAGCGCTTGA 21126  
QY 1141 TCAGGACATTTAGTGGGCTCATGACTCCGCGCAATTTGGCTGAACCGTTTCGGGAGG 1200  
DB 21127 TCAGGACATTTAGTGGGCTCATGACTCCGCGCAATTTGGCTGAACCGTTTCGGGAGG 21186  
QY 1201 GAGTCTTCTCTTCTGAGGATAGAGGATCTTCTGGGCAACATCTCTGGTGTGT-AA 1259  
DB 21187 GAGTCTTCTCTTCTGAGGATAGAGGATCTTCTGGGCAACATCTCTGGTGTGT-AA 21246  
QY 1260 ACATGAGTTTCTCTCATGATTTGGGACCTTATTCAGCTAAATGGGAATAGATCTGGCGG 1319  
DB 21247 ACATGAGTTTCTCTCATGATTTGGGACCTTATTCAGCTAAATGGGAATAGATCTGGCGG 21305  
QY 1320 GTTTGGTCCAGTTTGAAGCGGCTTATTTAGATCTCTCGCTGGTGTATTCAGCGGATA 1379  
DB 21306 GTTTGGTCCAGTTTGAAGCGGCTTATTTAGATCTCTCGCTGGTGTATTCAGCGGATA 21365  
QY 1380 TGAAGAAATCAGCGGATGTGCTTGAAGGATCTCAGAACTTCCAGCTGCGATCGCATC 1439  
DB 21366 TGAAGAAATCAGCGGATGTGCTTGAAGGATCTCAGAACTTCCAGCTGCGATCGCATC 21425  
QY 1440 TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCGCATGTTCAAGTCAAGCGGAT 1499  
DB 21426 TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCGCATGTTCAAGTCAAGCGGAT 21485  
QY 1500 TCAAGGAAAGAAACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 1559  
DB 21486 TCAAGGAAAGAAACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 21545  
QY 1560 TAAAGTGTGTACATCTGCACTCGCAATATTCAGCTGAGGATTTGCTGACATGCGA 1619  
DB 21546 TAAAGTGTGTGTACATCTGCACTCGCAATATTCAGCTGAGGATTTGCTGACATGCGA 21605  
QY 1620 TACCAATATTTTCAAGGACAGTGAACCAAGCGGTTGATTAAGCCATATGAAGGATC 1679  
DB 21606 TACCAATATTTTCAAGGACAGTGAACCAAGCGGTTGATTAAGCCATATGAAGGATC 21665  
QY 1680 GTCAAAACTTTTCTCTGATGACTGAAGCAAAATTTCTGTGTAGACATTTCTCCGCTTTG 1739  
DB 21666 GTCAAAACTTTTCTCTGATGACTGAAGCAAAATTTGTGTTAGACATTTCTCCGCTTTG 21725  
QY 1740 TGTCTCATGAGCGGATTCGCAAAAGCAGTGTATTGCTGAGCTATGAGTGCAGAGATCC 1799  
DB 21726 TGTCTCATGAGCGGATTCGCAAAAGCAGTGTATTGCTGAGCTATGAGTGCAGAGATCC 21785  
QY 1800 GA 1801  
DB 21786 GA 21787

RESULT 12  
AB025110 4660 bp DNA linear BCT 27-MAR-1999  
LOCUS AB025110  
DEFINITION Agrobacterium tumefaciens gene for indole acetamide hydrolase and

ACCESSION  
AB025110  
VERSION  
GI:4586311  
KEYWORDS  
tryptophan monoxygenase; indole acetamide hydrolase.  
SOURCE  
Agrobacterium tumefaciens (Rhizobium radiobacter)  
ORGANISM  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
REFERENCE  
1 (bases 1 to 4660)  
Kasahara, T., Endo, S., Ebinuma, H., Sugita, K., Kawaoka, A. and  
Moriwae, N.  
TITLE  
Agrobacterium tumefaciens plasmid PO22 indole acetamide hydrolase  
(iaah), tryptophan monoxygenase (iaaw) genes  
Published Only in Database (1999)  
JOURNAL  
2 (bases 1 to 4660)  
Kasahara, T., Endo, S., Ebinuma, H., Sugita, K., Kawaoka, A. and  
Moriwae, N.  
AUTHORS  
Direct Submission  
Submitted (18-MAR-1999) Takehide Kasahara, Nippon Paper Industries  
Co., Ltd. R&D Div., Wood Bio-engineering Central Research Lab.,  
5-21-1, Oji, Kita-ku, Tokyo 114-0002, Japan  
[E-mail: takeh@ntj.biglobe.ne.jp, Tel: 81-3-3911-5106 (ex. 292),  
Fax: 81-3-3914-3350]  
FEATURES  
source  
1..4660  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/strain="PO22"  
/db\_xref="taxon:358"  
complement (353..1756)  
/gene="iaah"  
complement (353..1756)  
/gene="iaaw"  
/codon\_start=1  
/transl\_table=1  
/product="indole acetamide hydrolase"  
/protein\_id="BAA76345.1"  
/db\_xref="GI:4586312"  
/translation="MYPTSLAQTLERLRKRYSCLEIVETLIARQAKPNNLLAT  
DWDGRLRSKAKDRHGNGLGCLPLCFKANIATGVFTSATPAILNHLPIKPSRV  
ABRLPSAGALPGASNMHLSFGITSNNYATGAVNPMPLIPGSSGVAAGVAVSR  
LMLGIGTGTGASVRLPALCGVGRFPLGIVYPRDRIIPSPTRDTAGIIAQCADV  
ILDDVYISGRSAKISPMPLKGRIGLPTTYFDLDDADVAFAETIRLLARGVTFF  
EADIVIELRNGSASLPALVPEPAKGLVDDPVAFSPVIGIRSPDVANIVSA  
QIDHQISNDEVELARQSRPRLOATRYRFFLYQIDALLPTAPLAARATIGESSVI  
HNGSMNTFKITVRNVDPSSNAGLPGLSLPACTLPURLPVGMELDGLASDRLLAIG  
MALEKAINSSFPDAFN"  
2096..4363  
/gene="iaaw"  
2096..4363  
/gene="iaam"  
/codon\_start=1  
/transl\_table=1  
/product="tryptophan monoxygenase"  
/protein\_id="BAA76346.1"  
/db\_xref="GI:4586313"  
/translation="WSASPLDNOCDHSTKAVDILMDKABELDRVSDASREAS  
RGRRTQISGECAGLACRLADGRPELSTSEKVAASAVIYVGEILGRILSESPW  
AARVSGVAIDLAPFCMDPSEAOULQTLFLSGRCASDSHFAVLSISTARST  
LQMPREKGTTRVGTFTLLEAVPDMVAAGRMLKASAGSPFTIDLIDYRSFF  
DCCSIRGIRGFPEDEVKPKVAIIGKISGLVAVASLHLAAGVDVITYASRVGKL  
WSHAPKDAVSVAEMGAMPAPASCLPEFLERIVLASSSMRPDPNCTUTNTIYOCIR  
YVWAKGQDPKPLFHRVYSGMRAPLADGRHEGIVLASPVITTOALSGDIRAHDSKQ  
TWLNPGRSFSSAIEIRIFLGTHPPEGSTWSPPHMDLCKLMIGISGGGFPVSESPF  
ILRLVINGYENQRMCSGISLSEPRRIASQVNVGSQRLRHVOVRAIIEKXTIK  
ILKSGISLSEYDKVVVTSGLANIQLRHCLCTDTTIFRAVNOAVNSHTSGSKFLFL  
TERKFMFLDIILPSVIMDGIKAVYCLDVEPDQNGKGLVPTTYMEDSHKLAVPD  
KKERFCLRDATSRSPAPAOHLVPACADYDQNVQHMILTENNAGGAPKTLNRGDF  
YSEELFPQALDPRNDTGYTLACSSCFPGWVEGAIQTRCANVCAI IHNCGSLAKDN  
FLHSMWKRYNNRN"

BASE COUNT 1174 a 1038 c 1171 g 1276 t 1 others  
ORIGIN

Query Match 88.7%; Score 1597.2; DB 1; Length 4660;  
 Best Local Similarity 93.6%; Pred. No. 0;  
 Matches 1687; Conservative 0; Mismatches 113; Indels 2; Gaps 2;

QY 1 ATGTCACCTTCACTCTCTCTGATTAACAGGCGATCATCTCCACCAAAATGATGAT 60  
 DB 2096 ATGTCACCTTCACTCTCTGATTAACAGGCGATCATCTCTACCAAAATGATGAT 2155

QY 61 CTGCAATGTCGATTAAGGCGGATGAAATGGAACCGGAGGTTTCCGATGCTTTGAA 120  
 DB 2156 CTGCAATGTCGATTAAGGCGGATGAAATGGAACCGGAGGTTTCCGATGCTTTGAA 2215

QY 121 CGAAGACCTTCTAGGGGAGAGGATTAATCAAAAGCTTCAACGAGTCAAGCGCTGAGT 180  
 DB 2216 CGTGAACCTTCTAGGGGAGAGGATTAATCAAAATCTCCGCGAGTCAAGCGCTGAGT 2275

QY 181 GCTTGGCAAAAGGCTGGCGGATGTCGCTTCCCGAGATCTCACTGATGAGTGAAGTGA 240  
 DB 2276 GCTTGGCAAAAGGCTGGCGGATGTCGCTTCCCGAGATCTCACTGATGAGTGAAGTGA 2315

QY 241 GTTCTCTCCGCTTATCTATATATGGAAGAAATTCGCGGCGGATTAATGAAATGAAA 300  
 DB 2336 GCTTCTCTCCGCTTATCTATATATGGAAGAAATTCGCGGCGGATTAATGAAATGAAA 2395

QY 301 CCTTGGCGGCGGCAAGAGTGAATGTCGCTTCCGATCACTGACACCATTTGATG 360  
 DB 2396 CCTTGGCGGCGGCAAGAGTGAATGTCGCTTCCGATCACTGACACCATTTGATG 2455

QY 361 GATTTCTCCGAAGACAATTAATCAAGCCCTGTTTGTCTGAGCGGTAAAGATGGA 420  
 DB 2456 GATTTCTCCGAAGACAATTAATCAAGCCCTGTTTGTCTGAGCGGTAAAGATGGA 2515

QY 421 CCGATGATCTTATGATCTGATGCGGCAATTCCTAAGATGCGCGCTTCAAGC 480  
 DB 2516 TCCAGCATCTTATGATCTGATGCGGCAATTCCTAAGATGCGCGCTTCAAGC 2535

QY 481 CTGCGCAATGCGCGCTGTAAGAAATGGAAGTGAATGCGGCTTCAAGC 540  
 DB 2576 CTGCGCAATGCGCGCTGTAAGAAATGGAAGTGAATGCGGCTTCAAGC 2635

QY 541 CTGGAAGGCGCGGCTGATTAAGATGGAATGGAATGGAATGGAATGGAATGGA 600  
 DB 2636 CTGGAAGGCGCGGCTGATTAAGATGGAATGGAATGGAATGGAATGGAATGGA 2695

QY 601 TCGGAGGATCTCTTCCCAACATGATGCTGTAAGAGTGAAGCGGTTTGAACAA 660  
 DB 2696 TCGGAGGATCTCTTCCCAACATGATGCTGTAAGAGTGAAGCGGTTTGAACAA 2755

QY 661 TGTTCGATAGTGAAGGATGCGCTTCTTCCGAGGATGCTTCAAGCGGAAATGAG 720  
 DB 2756 TGTTCGATAGTGAAGGATGCGCTTCTTCCGAGGATGCTTCAAGCGGAAATGAG 2815

QY 721 GTCAATGCGGCTGATTTCCGAGCTGCTGAGGCAACGATGCTTCAATGCTGAGGTA 780  
 DB 2816 ATCAATGCGGCTGATTTCCGAGCTGCTGAGGCAACGATGCTTCAATGCTGAGGTA 2875

QY 781 GAGCATGTTCAATATATGAAGCAAGTATGCTTGAAGGCAAGCTTTGGTCAATGCT 840  
 DB 2876 GAGCATGTTCAATATATGAAGCAAGTATGCTTGAAGGCAAGCTTTGGTCAATGCT 2935

QY 841 TTCAAGGAGCTCTAGTGTGTCGCGAAATGAGGCGGATGCGATTTCTCTGCTGCA 900  
 DB 2936 TTCAAGGAGCTCTAGTGTGTCGCGAAATGAGGCGGATGCGATTTCTCTGCTGCA 2995

QY 901 TTCTGCTGTTTTTTCTCTGAGCGGTTACGCGCTGCTTCTGATGAGCGGTTCCCAAT 960  
 DB 2996 TTCTGCTGTTTTTTCTCTGAGCGGTTACGCGCTGCTTCTGATGAGCGGTTCCCAAT 3055

QY 961 CCGGCGCAGTTCGACATTAATCTTGAAGGCGTCAAAATCATGAGGAAAGCGG 1020  
 DB 3056 CCGGCGCAGTTCGACATTAATCTTGAAGGCGTCAAAATCATGAGGAAAGCGG 3115

QY 1021 CAGCTGCACCGAAGCTGTTCAATCGGCTTACAAACGTTGGCGCTTCTTGAAGAC 1080

DB 3116 CACGACCGACCGAAGCTGTTCCATCCGTTTACAGGGGTTGGGCTGTTCTGAGGAGAC 3175  
 QY 1081 GGTTCATGAGGAGATATGTTGTTGGCTTGGCTGCTGCTATTAATCTGAGGCTTGAA 1140  
 DB 3176 GGTTCATGAGGAGATATGTTGTTGGCTTGGCTGCTGCTATTAATCTGAGGCTTGAA 3235

QY 1141 TCAGGACATTAATGAGGAGCTCATGATCTCTGCAAAATTTGCTGAACCGTTTCGGAG 1200  
 DB 3236 TCAGGAGACATTAATGAGGAGCTCATGATCTCTGCAAAATTTGCTGAACCGTTTCGGAG 3295

QY 1201 GAGTCTTCTCTTCAAGGATAGAGGATCTTTCTGAGGACACATCTCTGTTGTT-AA 1259  
 DB 3296 GAGTCTTCTCTTCAAGGATAGAGGATCTTTCTGAGGAGATCTTCTGTTGTTGAA 3355

QY 1260 ACATGAGTTTCTCATGATTTGGGACCTATTAAGCTATAGGAAATGAGATCTGGCGG 1319  
 DB 3356 ACATGAGTTTCTCATGATTTGGGACCTATTAAGCTATAGGAAATGAGATCTGGCGG 3414

QY 1320 GTTGGTCCAGTTTGTGAAAGCGGGTTTATTAAGATCTTCCGCTGATCAACGATA 1379  
 DB 3415 GTTGGTCCAGTTTGTGAAAGCGGGTTTATTAAGATCTTCCGCTGATCAACGATA 3474

QY 1380 TGAAGAAATACGCGATGTCCTTGAAGGATCTCAGAACTTCCAGTGGATTCGATC 1439  
 DB 3475 TGAAGAAATACGCGATGTCCTTGAAGGATCTCAGAACTTCCAGTGGATTCGATC 3534

QY 1440 TGAAGTTTAAAGGATGTCGTCGTCGAGCGGATGTCATGTCATGTCAGGCGGAT 1499  
 DB 3535 TGAAGTTTAAAGGATGTCGTCGTCGAGCGGATGTCATGTCATGTCAGGCGGAT 3594

QY 1500 TCAGAGGAAAGACAAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 1559  
 DB 3595 TCAGAGGAAAGACAAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGA 3654

QY 1560 TAAAGTGTGTGTCATCTGAGTCCGAAATATCAACTCAGGATTTGCTGATGCGA 1619  
 DB 3655 TAAAGTGTGTGTCATCTGAGTCCGAAATATCAACTCAGGATTTGCTGATGCGA 3714

QY 1620 TACCAATTTTTTCAAGGACCGAGTGAACCAAGGGTGAATTAAGCATATGACAGAT 1679  
 DB 3715 TACCAATTTTTTCAAGGACCGAGTGAACCAAGGGTGAATTAAGCATATGACAGAT 3774

QY 1680 GTCAAAATCTTCTCTATGATGATGAAGAAATTTCTGTTAGACATATCTCCGCTT 1739  
 DB 3775 GTCAAAATCTTCTCTATGATGATGAAGAAATTTCTGTTAGACATATCTCCGCTT 3834

QY 1740 TGTCTCATGAGCGGATGCGAAAGCAGTGTATGCTGAGTATGAGTCCGAGATCC 1799  
 DB 3835 TGTCTCATGAGCGGATGCGAAAGCAGTGTATGCTGAGTATGAGTCCGAGATCC 3894

QY 1800 GA 1801  
 DB 3895 GA 3896

RESULT 13  
 ATTAAAH  
 LOCUS  
 DEFINITION  
 A. tumefaciens Tm4 Ti plasmid DNA with TA-iaah interrupted by IS866,  
 TA-iaah, T-1pt and T-6 genes.  
 ACCESSION  
 X56185.1 GI:39133  
 VERSION  
 KEYWORDS  
 iaah gene; iaah gene; indoleacetamide hydrolase; 1pt gene;  
 isopentenyl transferase; tyriophane monooxygenase.  
 SOURCE  
 Agrobacterium tumefaciens (Rhizobium radiobacter)  
 ORGANISM  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 1 (bases 7402 to 8783)  
 REFERENCE  
 Bonnard, G., Thirland, B., Paulus, F., Szegedi, E. and Otten, L.  
 Nucleotide sequence, evolutionary origin and biological role of a  
 rearranged cytokinin gene isolated from a wide host range biotype

III Agrobacterium strain  
JOURNAL Mol. Gen. Genet. 216 (2-3), 428-438 (1989)  
MEDLINE 89313683  
PUBMED 2546041  
REFERENCE  
2 Bonnard, G., Vincent, F. and Otten, L.  
TITLE Sequence of Agrobacterium tumefaciens biotype III auxin genes  
JOURNAL Plant Mol. Biol. 16 (4), 733-738 (1991)  
MEDLINE 91329707  
PUBMED 1868204  
REFERENCE  
3 (bases 1 to 10200)  
Bonnard, G.  
TITLE Direct Submission  
SUBMITTED (03-DEC-1990) G. Bonnard, INSTITUT DE BIOLOGIE  
MOLECULAIRE DES PLANTES DU CNRS, 12 RUE DU GENERAL ZIMMER, 67084  
STRASBOURG CEDEX, FRANCE  
FEATURES  
source Location/Qualifiers  
1. 10200  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/strain="Tm4"  
/db\_xref="taxon:358"  
/plasmid="T1"  
/complement(join(192..653,3369..4319))  
/gene="TA-1aah"  
complement(join(192..653,3369..4319))  
/gene="TA-1aah"  
/codon\_start=1  
/transl\_table=11  
/product="indoleacetamide hydrolase"  
/protein\_id="CAA39649.1"  
/db\_xref="GI:39137"  
/db\_xref="SWISS-PROT:P25016"  
/translation="MPTSLAQTLELRKDYSCLELYETLLIARCEAKALNLIAT  
DWGLRRRAKIDIRHNGAVGLRGIPLCEPANTATGI PPTATATPALINLEPIERI  
AERLFSAGALPGASGMWHELSPGITSNNVTGAVRNPMWLSLPGSSGVAVAASR  
LMLGIGTGTGASVRLPALCGVAGRPRTGAPRRIIPVSPRTAGIACQADV  
VLIDOVISGRPARILPILKRLRIGLPTTFYVDLDADVAFAAETTRILANGVFEV  
EADPIHLEDNSGASLPALXERPHALKQYLDPPVCTVSGSSSDVIKEIRSDVANI  
VNAQIDGHQTSNAEYELAROSFRPLQAAATNFRILRLDAIILFPAPLAAKIGQDS  
SVHNGSMVNTFKIYVAVNDPSSNAGLPGSLPVLGTLPRLLPVMEIDIGLAGDHRLL  
ALGALEKAINFRSPDVNL"  
646..653  
654..3369  
/insertion\_seq="IS866"  
3370..3377  
4485..4490  
/note="Sep I junction site"  
4657..6924  
/gene="TA-1aam"  
4657..6924  
/gene="TA-1aam"  
/transl\_table=11  
/product="tryptophane monooxygenase"  
/protein\_id="CAA39646.1"  
/db\_xref="GI:39134"  
/db\_xref="SWISS-PROT:P25017"  
/translation="MSASSLDKQCHDFSTKIVLDIMVDAKELDRVAAAFSEBEAS  
REBRIQISGEENAGLAKRLADGRPEISAGORVAVLSAYIVGEISLRMLIEPAS  
VTRRVSGVAILDAPSCMDISRAOLQTMMLLSGKXCAPDLSFPVAISISETARST  
LQAPYBEGSLKSTGPTVYIEAVPDMVAYRNMLKASNSPTTDLIDYRFL  
DKDSBGRITFPEDVPRFVAVTGAIGISGLVAVSELHAGVDVTTIEGDLVGGKL  
WSHAFDAPGVAVEMAMRP PPAASLGFLELRYGLSMPFPENPGTVTDLYBCKR  
YMMKAOQPEPKLFHRYVSGWHAFLKDGFLGDIIVLASPDALTKALSGDI RRAMDWQ  
IIMNRGREGFSIAIERIFLGTHPGSETWSFPHMDLFLMGLISGGSGPVESEGFT  
EILRLVINGYEENOMCSGISELPRISAOVNVGVSQRIADVAVOVNRAIIEKXKIX  
ILKSGISELYDKVVVTSGLANTOLRHCILCTTIRAPVNOAINDSHMTSGSKFLI  
TEKFWFDHMLPESCVMIDGPAKAVYCLIDIEFPDPNGKGLVLIYTTEDDSHKLAVPD  
KSERLCILRDALDKISFPVEFAQHLVPACADIDQNVQHDWLTDENAGREFLNKRGDF  
YSEELFFQALDITNDTGVYLAGSCSFTGGWVGAIQTACNAVCAIITHNCGGILLADN  
PLGHPMKRYVYRNRN"  
7613..8335  
gene

CDS  
/gene="T-1pc"  
7613..8335  
/gene="T-1pc"  
/codon\_start=1  
/transl\_table=11  
/product="isopenicillin transferase"  
/protein\_id="CAA39647.1"  
/db\_xref="GI:39135"  
/db\_xref="SWISS-PROT:P25018"  
/translation="MDRLIRPPTGKSTALAAQQTGLPYLSIDRYQCCPOLSTG  
SGRPVVEELKTRLYLDDRLVKGI IAEQAHKRLIAEVYEAHAGLILGSGISL  
LKCMQSGYWSADFRMHIIRKLADBEETFMKAKAVKQMLCPAIGPSLIQGLVLYLM  
EPLRLPILKEIDGYRYVAMLFASQNRITPDMILQLDADMEGKLHIGIAQYLLIAROE  
HEFPVSAAPFEGPPEGAY"  
complement(9013..9636)  
/gene="T-6b"  
/gene="T-6b"  
complement(9013..9636)  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA39648.1"  
/db\_xref="GI:39136"  
/db\_xref="SWISS-PROT:P25019"  
/translation="MTVANMQRDEFTRLINAGELQGRLEQARTDREALIAEIVYRHP  
GATPEGDEYILTCGLVYVYLSQTEARCLNBLPSNSNFGVTVTAIPPMWMDT  
OTNLTLQRCDDGGIVNYHGSRTGFLATMLNCFVRFGTDEINGASVGFYARAG  
NYTERGEDDNDNIGDEGAGAEIRDYQGDLYNVPYVALSSRLSA"  
BASE COUNT 2436 a 2471 c 2709 g 2584 t  
ORIGIN  
Query Match 81.5%; Score 1467.6; DB 1; Length 10200;  
Best local Similarity 89.1%; Pred. No. 0;  
Matches 1606; Conservative 0; Mismatches 194; Indels 2; Gaps 2;  
QY 1 ATGTGAGCTTCACTCTCTTGATATAACAGTCGATCATCTCCAAACAAATGGTGAT 60  
Db 4657 ATGTGAGCTTCACTCTCTTGATATAACAGTCGATCATCTCCAAATGGTGAT 4716  
QY 61 CTGACAAATGCTGATTAAGCGGATGAATTGACCCGAGGCTTCCGATCCTTTAGAA 120  
Db 4717 CTGATATAGTGTGACAAAGCTGATGATGACCGCGGTTGCCACCTTCTCAGAA 4776  
QY 121 CGAGAAAGCTTCAAGGGAAGAGATTAATCAAGCTCCAGAGTGCAGCGTGAGTTA 180  
Db 4777 CGAGAAAGCTTCAAGGGAAGAGATTAATCAAGCTCCAGAGTGCAGCGTGAGTTA 4836  
QY 181 GCTTGCAAAAGCTGCGCGATGTCCTTCCCGAGATCTCAGCTGTGGAAGATGACA 240  
Db 4837 GCTTGCAAAAGCTGCGCGATGTCCTTCCCGAGATCTCAGCTGTGGAAGATGACA 4896  
QY 241 GTTCTCTCGCTTATATCTATATTGGCAAGAAATTCTGGGCGGATTAATGATGAAA 300  
Db 4897 GTTCTCTCGCTTATATCTATATTGGCGAGAAATTCGAGGTGATATCTCAACACAGA 4956  
QY 301 CCTGGGCGCGGCAACAGTGAAGTGTCTGTGCATGATGATGGAACCAATTTGCATG 360  
Db 4957 GCTTCGCTGCAACAGTGAAGTGTCTGTGCATGATGATGGAACCAATTTGCATG 5016  
QY 361 GATTTCTCCGAAGCAACAATTAATCAAGCGCTGTTTTCTGAGCGGTAAAGATGTGA 420  
Db 5017 GATTTCTCCGAAGCAACAATTAATCAAGCGCTGTTTTCTGAGCGGTAAAGATGTGA 5076  
QY 421 CCGATTGATCTTAATGATTTGTCGTCATTTCAATCTTAAGACTGCGCGCTTTCGAAC 480  
Db 5077 CCGAGGACCTTAATGATTTGTCGTCATTTCAATCTTAAGACTGCGCGCTTTCGAAC 5136  
QY 481 CTGCAATGCGCTGTGACAGAAATGCGAGTAAGCGGTTTTCATTAAC 540  
Db 5137 CTGCAATGCGCTGTGACAGAAATGCGGTTTTCATTAAC 5196  
QY 541 CTGGAAGGGCGGTGACATTTGACATGTGACTTATGTGCAAACTGATGTGAAGGT 600  
Db 5197 ATTGAAGAGGACGTAACATTTGACATGTGACTTATGTGCAAACTGATGTGAAGGC 5256

QY 601 TCGGAGAGTTCCTTCCAAACATCGACTTGCTACAGCTGACAGACCGTTTTTTGACCAA 660  
DB 5257 TCGGAGAGGTCCTTCCAAACATCGACTTGCTACAGCTGACAGATTTGTTCTCGACAAA 5316  
QY 661 TGTTCGATAGTGAAGCGATCGGCTTCTTCCGAGAGATTTCTTAAGCCGAAAGTGGCG 720  
DB 5317 TGTTCGATAGTGAAGCGATCGGCTTCTTCCGAGAGATTTCTTAAGCCGAAAGTGGCG 5376  
QY 721 GTGATTGGCGCTGGCATTTCCGAGCTGCTGAGTGGCAAAACGATCGCTTCACTGCTGGGTA 780  
DB 5377 GTGATTGGCGCTGGCATTTCCGAGCTGCTGAGTGGCAAAACGATCGCTTCACTGCTGGGTA 5436  
QY 781 GACGATGTTCAATATATGAAAGCAAGTATCGTGTGAGGCAAGCTTTGTCACATGCT 840  
DB 5437 GACGATGTTCAATATATGAAAGCAAGTATCGGTTGAGAGTAAGCTTTGTCACATGCC 5496  
QY 841 TTCAGGAGCGCTCTAGTGTCTGAGCGGAAATGGGGCGATGCGATTCTCTGCTGCA 900  
DB 5497 TTCAGGAGCGCTCTAGTGTCTGAGCGGAAATGGGGCGATGCGATTCTCTGCTGCA 5556  
QY 901 TTCTGCTTTGTTTTTCTCTCTGAGCGTTACGCGCTGTCTTGCATGAGCGCGTTCCCAAT 960  
DB 5557 TTCTGCTTTGTTTTTCTCTCTGAGCGTTACGCGCTGTCTTGCATGAGCGCGTTCCCAAT 5616  
QY 961 CCGGCGACAGTCGACCTTACCTTACCAAGCGCTCCAAATCATGTGAAAGCCGGG 1020  
DB 5617 CCGGCGACAGTCGACCTTACCTTACCAAGCGCTCCAAATCATGTGAAAGCCGGG 5676  
QY 1021 CAGCTGCACCGAAGCTGTTCCATCGGCTTACCAAGCGTTGGCGTGGCTTTGAAAGAC 1080  
DB 5677 CAGAGCGACCGAAGCTGTTCCATCGGCTTACCAAGCGGTTGGCGTGGCTTTGAAAGAC 5736  
QY 1081 GGTGTTTATGAGCAGATATTGTGTGTGGCTTCCGCTGTCTTACTACAGCGCTTAAA 1140  
DB 5737 GGTGTTTATGAGCAGATATTGTGTGTGGCTTCCGCTGTCTTACTACAGCGCTTAAA 5796  
QY 1141 TCGAGCACATTTAGTGGGCTCATGATCTCCGCGCAATTTGGCTGAAACCGTTTGGGAGG 1200  
DB 5797 TCGAGGACATTTAGGCGGCTCATGATCTCCGCGCAATTTGGCTGAAACCGTTTGGGAGG 5856  
QY 1201 GAGTCTCTCTCTTCAGGATAGAGAGATCTTCTGCGGCAACATCTCTGTGTG-TAA 1259  
DB 5857 GAGTCTCTCTCTTCAGGATAGAGAGATCTTCTGCGGCAACATCTCTGTGTGAGAA 5916  
QY 1260 ACATGAGTTTCTCTCATGATGAGACCTTATCAAGCTAATGGAAATAGATCTGGGGG 1319  
DB 5917 ACATGAGTTTCTCTCATGATGAGACCTTATCAAGCTAATGGAAATAGATCTGGGC-GG 5975  
QY 1320 GTTTGTCAGTTTGTAAAGCGGTTTATGAGATCTCCGCTTGGCTATCAACGAGTA 1379  
DB 5976 GTTTGTCAGTTTGTAAAGCGGTTTATGAGATCTCCGCTTGGCTATCAACGAGTA 6035  
QY 1380 TGAAGAAATCAGCGGATGTGCCCTGAAGGAATCTCAGAACTTCCACGTCGGATCGCATC 1439  
DB 6036 TGAAGAAATCAGCGGATGTGCCCTGAAGGAATCTCAGAACTTCCACGTCGGATCGCATC 6095  
QY 1440 TGAAGTGTAAAGGTGTGTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGGAT 1499  
DB 6096 TCAAGTGTAAAGGTGTGTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGGGAT 6155  
QY 1500 TCGAAGGAAAGCAAAATTAAGATTAAGCTTAAGGCGGAGTATCTGAACCTTTATGA 1559  
DB 6156 TCGAAGGAAAGCAAAATTAAGATTAAGCTTAAGGCGGAGTATCTGAACCTTTATGA 6215  
QY 1560 TAAAGTGTGTCAATCTGGAAGTCGCAAAATATCAATCTCAGCATTTGCTGACATCGA 1619  
DB 6216 TAAAGTGTGTCAATCTGGAAGTCGCAAAATATCAATCTCAGCATTTGCTGACATCGA 6275  
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATACAGCCATATGACAGATC 1679  
DB 6276 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATACAGCCATATGACAGATC 6335

QY 1680 GTCAAAATCTTCTCGATGACTGAAGCAAAATTTCTGTTAGACATATCTCCGTTCTTG 1739  
DB 6336 TTCAAAATCTTCTCGATGACTGAAGCAAAATTTGTTGACATATGCTCCGTTCTTG 6395  
QY 1740 TGTTCCTCATGAGCGGATGCAAAAGCAATGTTTCCCTGAGTATGATGCGAGATCC 1799  
DB 6396 TGTTCCTCATGAGCGGATGCAAAAGCAATGTTTCCCTGAGTATGATGCGAGATCC 6455  
QY 1800 GA 1801  
DB 6456 GA 6457  
RESULT 14  
AVU83987  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
source  
source  
misc\_feature  
repeat\_region  
CDS  
gene  
15463 bp DNA linear BCT 02-MAR-2001  
Agrobacterium vitis plasmid pTiTm4 agrocinopine synthase and (5)  
genes, (laah) pseudogene, and (laah), (lpc), (6b) and (oce) genes,  
complete cds.  
U83987.1 GI:1814322  
Agrobacterium vitis (Rhizobium vitis)  
Agrobacterium vitis  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
Osten, L. and De Ruffray, P.  
Sequence and functional analysis of the shoot-inducing octopline  
T-DNA of Agrobacterium vitis strain C6474  
Unpublished  
Osten, L. and De Ruffray, P.  
Direct Submision  
Submitted (07-JAN-1997) Institute for Plant Molecular Biology, rue  
du General Zimmer 12, Strasbourg 67084, France  
Location/Qualifiers  
1. 15463  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/strain="Tm4"  
/db\_xref="taxon:373"  
/plasmid="pTiTm4"  
3657. 6588  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:373"  
/insertion\_seq="I5866"  
15438. 215463  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:373"  
/insertion\_seq="I5867"  
139. 15310  
/note="shoot-inducing octopline T-DNA"  
139. 165  
/note="flanking repeat; left border of the T-DNA"  
complement(373. 1629)  
/codon\_start=1  
/transl\_table=11  
/product="agrocinopine synthase"  
/protein\_id="AAB41872.1"  
/db\_xref="GI:1814323"  
/translation="MEGTAPMAPMTLHMDLPEVTPPEVLARHVPKLEBELLEPLQ  
VIEHGFNDGNRIQECTASLLAALGGGRNLSLDVCLTSNITVVSDDLTWYS  
EKGDGFENEIHSDDINNVPIIIRVSNGLIDRVETDIDPLDELKIVSNPD  
ATIFLDGRNVEAIVAWLSHREYQRAVLYFYREYPDGAFVDAVLPAQASMR  
KSIAPLAPPEBELCRLARQVSEPTVDLYAGKAWIDSLMDMRVVAIVVSE  
VSRNLDIVYDKVILAFPSDOAAVRLAHYKEDTMIKARPLKFAAVTRCYDFPAL  
LNSGEAERSIDINTGARPHETDERKHIRMRKGTGNSATIDAWIISDSBEDMAIW  
EMRNGIDREVSHLSPHLDVDTSE"  
2235. 2918  
/gene="5"

```

CDS      2235..2918
         /gene="5"
         /codon_start=1
         /transl_table=1
         /protein_id="AAB41873.1"
         /db_xref="GI:1814324"
         /translation="MDHSRPPIPIVDCSHIODRELEIVLATRRATIRPAQODLIRA
         QNSWIDSTITDINVPIDPAIDEVKRRFRVACLPAPVGKSLIVLINDSLMYVCSFQSM
         RRYACERPEFGSGKGVIVSTVPSPQITKEAMQWNEVCHITNNEKDLDIAYF
         LNTSLQNTSPFSHVXIGDPAFLAPSRADPFCVEIVAYGKALFQDNGLKAPKAGWSMA
         LTSLLKRLV"
         complement(3403..7530)
         /gene="1aah"
         complement(join(3403..3856,6589..7530))
         /gene="1aah"
         /note="interrupted by the bacterial insertion element
         1S866"
         /pseudo
         /codon_start=1
         /transl_table=1
         3857..3864
         /note="flanking repeat; left border of the 1S866 insertion
         element"
         /rpt_type=direct
         6581..6588
         /note="flanking repeat; right border of the 1S866
         insertion element"
         /rpt_type=direct
         7868..10135
         /gene="1aam"
         7868..10135
         /gene="1aam"
         /codon_start=1
         /transl_table=1
         /protein_id="AAB41874.1"
         /db_xref="GI:1814325"
         /translation="WSASSLDDKQDHFSTYIVDLINWKADELDRVAAAFSEKAS
         RRRRIQSJSGECNAGLACKRLADGRPEPISAGORVAVLSAYIVGEBILPMILPEKAS
         VTRRSGVAIDLAPSCMDISRAQLQTMNLISGRCAPSDLSHFVAISISTERSRT
         LOMAYREBSLKSVTGFTVITIEAVPDMVAVGNRLMKASAGSPPTDLLDYRFL
         DKCSGSRGKFPEDVPFRKAVVIGAGISGLVVASSELLHAGVDVTVYENDGRVCGKL
         WSHAKDAPGVAAEKGAMRFPFPAASCLFFPLEIRGLSMPFPNPGVTVDVLYVEGR
         IWMKGGQOPPLKFRVYVSGWHAFLKDGLESDIVLASPDALTEALKSGDILRADDSQ
         IWLNFGRGSESSAIERIFLGTHPPGGETWSFPHWDLFKLMGIGSGGFGVPESSGPT
         ELIRIVINGEENOMCSGEGISLEPRLASOVNGSVSORIRHVOYRAIEKEKTKIK
         IRLKGISLKYKVVVTSGLANIORHCLTCDTTIPRAPVQAVDNSHMTGSSKFLPL
         TERKCMPEHMLPSCVLMDSGFKAKAYCUDYEPQDNRGKGLVISTTWEDESHKLAVD
         KKERCLMCLDAISKSPVPAQKLVPAACADYQNVAVQHMWLTDENAGRFKLNREGDF
         YSEELFQALDITDNTGYVLAGSCSFGTWEGAIQYACNAVCAIITHNGQILAKON
         PLKHFWKYNYNENR"
         10824..11546
         /gene="1pc"
         10824..11546
         /gene="1pc"
         /codon_start=1
         /transl_table=1
         /protein_id="AAB41875.1"
         /db_xref="GI:1814326"
         /translation="WDLRLIRGPTCTGKTSATATAAOQTGIPVLSLRVQCCPOLSTG
         GSRLPVEELKGTSLRYLDDRLPVKGIIAEQAHERLLAEVNYFAHGGLILLEGSSISL
         LKCMQSGWMSADFERFWHIIIRKLDADEETFMMAAARVQMLCPAIGSLIDELVYLN
         BRRLLRPIIAKEIDIGRYVAMLPAFSONRITPDMLLQDADMEGLHIGIAOEYVLIHAROE
         HEFPVPSAAKEIFGFGPEGPFPGAY"
         complement(12224..12847)
         /gene="6b"
         complement(12224..12847)
         /gene="6b"
         /codon_start=1
         /transl_table=1
         /protein_id="AAB41876.1"
         /db_xref="GI:1814327"
         /translation="WTVANMOVDRFTRLINAGELQGRLEQARLTDPGALLAEIYFHPD
         GATPEAGDEYILLTQOGLVYVYLSEQTRQCALNRLLPSSNSNFGTIVVTAIPMLMDT

```

gene	CDS	repeat_region	BASE COUNT	3873 a	3665 c	3902 g	4023 t	ORIGIN
gene		Query Match	81.5%	Score 1467.6;	DB 1;	Length 15463;		
		Best Local Similarity	89.1%	Pred. No. 0;				
		Matches 1606;	Conservative 0;	Mismatched 194;	Indels 2;	Gaps 2;		
		1	ATGTCACCTTCACCTCTCTCTGATTAACAGTCGCATATCTCCCAACCAAAATGGTGAT	60				QTALNTLQBERCDGGIVNYHHGSTRNFPLAIALMISNCFVRGTDEINGASTGFYARRK
		7668	ATGTCACCTTCATCTCTCTGATTAACAAATGCGATCTATTTCTTACCAAAATGTGGAT	7927				NTYEEGDDDDNEIDEEAGAEITRDQFDLVNYPVIALGSSRLA"
		61	CTGACAAATGTCGATTAAGCGGATGAATTGGACCGCAGGGTTCCGATGCTTTTAGAA	120				complement (131779..14855)
		7928	CTGATTAATGTCGCAAGGCTGATGAATGGACCGCGCGTTCGCCGACCTTCTCGAA	7987				/gene="ocs"
		121	CGAAGACCTTCTAGGGGAGAGAGATTAATCTCAAGCTCCACCGAGTCAAGCGCTGGGTTA	180				/codon_start=1
		7988	CGAAGACCTTCTGAGGAGAGAGAAAGAAATTAATGTCAAATTTCCGCGAGTCAAGCGCTGGGTTA	8047				/transl_table=1
		181	GCTTGCAGAAAGGCTGGCGGATGTCGCTCCCGCAGATCTCAGCTGGTGAAAGTATGCA	240				/protein_id="AAB41877.1"
		8048	GCTTGCAGAAAGGCTGGCGGATGTCGCTCCCGCAGATCTCAGCTGGTGAAAGTATGCA	8107				/db_xref="GI:18143328"
		241	GTTTCTCTCCGCTTAATCTAATTTGGCAAGAAATTTGGGGCGGATACTTTGAATCGAA	300				/translation="MAKVALILGAGNLLTLAAGLARRLGGTSPIMADPLSNSSFNVDYR
		8108	GTTTCTCTCCGCTTAATCTAATTTGGCAAGAAATTTGGGGCGGATACTTTGAATCGAA	8167				CLGSEIVGPDYDGDPOPLRDLDTGATISGAAPLPLVTPVQGGILRELAKPULSNS
		301	CTTTGGGCGGGGGGCAACAGTGAATGTCGTCTGTTGCCATCGACTTGGACACATTGGCATG	360				VVALPSATSLACKQTLVPTFPAPAVIBESTSPYACRRKARVLMILGVAKFEVATTT
		8168	GCTTCGGGTGGCAACAGTGAATGTCGTCTGTTGCCATCGACTTGGACACATTGGCATG	8227				QPISEEVKGRFVELFPNPQWYOHIPASIFSNTPVAPHPGILAAIRISIEGILFVPP
		361	GATTTCCTCCGAAGCAACAATTCACAAGCCCTGTTTGGCGAGCGGGTAAAGATGTGCA	420				FYRFPQVQATIRVLAIDBERLSIVDAIGLSEITDFYSKWKYGGHACNAEFTYEGE
		8228	GATTTCCTCCGAAGCAACAATTCACAAGCCCTGTTTGGCGAGCGGGTAAAGATGTGCA	8287				YAEIETPKMNHRLTDEDVKHILVLMWEIABAIQVQPEMKSVVQEASDVLENILRLT
		421	CCGATTGATCTTAGTCACTTTCGTGGCCATTTCATCTCTAAGACTGCGCGCTTTGCAAC	480				GRGISLNLGGANNAIVRALNGV"
		8288	CCGATTGATCTTAGTCACTTTCGTGGCCATTTCATCTCTAAGACTGCGCGCTTTGCAAC	8347				15287..15310
		481	CTGCGCAATGCCGCTGTACGAGATGCAACGATGAATGCGTTACCGGGTTTACATTAAC	540				/note="flanking repeat; right border of the T-DNA"
		8348	CTGCGCAATGCCGCTGTACGAGATGCAACGATGAATGCGTTACCGGGTTTACATTAAC	8407				
		541	CTTGAAGGGGGCGGCGCACTTTCAGATGATGCTTAATGCTGGAACCTGATGCTGAAGGGT	600				
		8408	ATTGAAGGGCGATGACATTTGACATGTATGCTTAATGCTGGAACCTGATGCTGAAGGGT	8467				
		601	TCGGCAGGTTCTTTCCAAACATTCGATCTCTACGACTGCAAGCCGTTTTCGACAA	660				
		8468	TCGGCAGGTTCTTTCCAAACATTCGATCTCTACGACTGCAAGCCGTTTTCGACAA	8527				
		661	TGTTTCGATAGTGAACGATCGGCTTTTTCGAGAGATGTTCTTAAGCCGAAAGTGGCG	720				
		8528	TGTTTCGATAGTGAACGATCGGCTTTTTCGAGAGATGTTCTTAAGCCGAAAGTGGCG	8587				

QY	721	GTCAATGGCGCTGGCATATTTCCGAACTCGTGTGGCAAAACGAACTCTTCAGCTGGGGTA	780
Dp	8588	GTCAATGGGGCTGGCATATTTCCGGGCTCGTGGTGGCAACGAACTCTTCAGCTGGGGTA	8647
QY	781	GACGATGTTACAAATATATGAAAGCAATGATCGTGTGGAGCAACTTTGGTCAATGCT	840
Dp	8648	GACATATGTTACAAATATATGAAAGCAAGTGATCGGGTTGAGGGTAAGCTTTGGTCAATGCTC	8707
QY	841	TTCAAGGACGCTCCCTAGATGCTGTGGCCGAAATGGGGGGGAGATGGCATTTTCCCTCGTGTCA	900
Dp	8708	TTCAAGGACGCTCCCGGGGCGTGTGGCCGAAATGGGGGGGAGATGGCATTTTCCCTCGTGTCA	8767
QY	901	TTCTGCTGTTTTTTCTTCTCTCGAGCGTTACGGCCTGTCTTCGATGAGGCCGTTCCCAAT	960
Dp	8768	TCGTGCTGTTTTTTCTTCTCTCGAGCGTAACGGTCTGTCTTCGATGAGGCCGTTCCCAAT	8827
QY	961	CCCGGCAACGTGCAACCTTACTTGGTCTACCAAGCGCTCCAAATCATGTGGAAAGCCGGG	1020
Dp	8828	CCCGGCAACGTGCAACCTGACTTGGTCTACAGAGGGTTGCCAATCATGTGGAAAGCCGGG	8887
QY	1021	CAGCTGCACCGAAGCTGTCCATCGCCTTTTACAACGGTTGGCGTGCCTTCTTGAAGAC	1080
Dp	8888	CAGCAGCCACCGAAGTTGTTCCATCGCCTTTTACAGCGGGTGGCATGCGTTCTTGAAGAC	8947
QY	1081	GCTTTTCATGAGCGAGATATTTGTGTGGCTTCGCTGTGCTATTACTCAGGCTTGAA	1140
Dp	8948	GCTTTTCCTTGAAGGAGATATTTGTGTGGCGTGCCTGTATGCTATCATCAGAGGCTTGAA	9007
QY	1141	TCAGAGCACTTAAGTGGGCTCATGACTCTGGGCAAAATTTGGCTGAAACGTTTTGGGAGG	1200
Dp	9008	TCAGGGGCACTTAAGCGGGCTCATGACTCTGGCAAAATTTGGCTGAAACCGTTTTGGGAGG	9067
QY	1201	GAGTCTTCTCTTCAGGGATAGAGAGATCTTCTGGGACACATCTCTGTGTGG- TAA	1259
Dp	9068	GAGTCTTCTCTCTCAGCGATAGAGAGATCTTCTGGGACACGATCTCTGTGTGGAGAA	9127
QY	1260	ACATGAGATTTTCCCATGATTTGGACCTATTCAAGCTAATGGGAATAGAACTGTGGCGGG	1319
Dp	9128	ACATGAGATTTTCCCATGATTTGGACCTATTCAAGCTAATGGGAATAGAACTGTGGC- GG	9186
QY	1320	GTTTGTGTCAGATTTTTTGAAGCGGGTTATTAGAGATCTCCGCTGTGATCAACGAGATA	1379
Dp	9187	GTTTGTGTCAGATTTTTTGAAGCGGGTTACTGAGATCTTGGCTTGTGCATTAAGCGATA	9246
QY	1380	TGAAGAAATCAGCGGATGTGCGCTGGAAGAACTCAGAACTTCCACGTGAGTGCATC	1439
Dp	9247	TGAAGAAATCAGCGGATGTGCTGGAAGAACTCAGAACTTCCACGTGCAATAGCTCTC	9306
QY	1440	TGAAGTGTTAACGGTGTGTCGTGAGGACAGGCAATAGCCATGTTCAAGTCAGGGCGAT	1499
Dp	9307	TCAAGTGTTAACGGTGTGTCGTGAGGACAGGCAATAGCCATGTTCAAGTCAGGGCGAT	9366
QY	1500	TCAGAGGAAAGACAAAAATTAAGATTAAGGCTTAAGACGGGATATCTGAACCTTAATGA	1559
Dp	9367	CGAAGAGAAAGACAAAAATTAAGATAAGGCTTAAGACGGGATATCAGAACTTTATGA	9426
QY	1560	TAAAGTGTGTGCATCTTGAATCCGCAAAATTCACAACTCAGGCACTTGTCTGACATGCGA	1619
Dp	9427	TAAAGTGTGTGTATCATCTGGAATCCGCAAAATTCACAACTCAGGCACTTGTCTGACATGCGA	9486
QY	1620	TACCAATATTTTTTCAGGACCAAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATC	1679
Dp	9487	TACCAATATTTTTTCAGGACCAAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATC	9546
QY	1680	GTCAAAATCTTTCTCTGATGACTGAACGAAATTTCTGTTAAGCATATCTCCGTCCTTG	1739
Dp	9547	TTCAAAAATCTTTCTCTGATGACTGAACGAAATTTTGGTTTGAACCATATGCTCCGTCCTG	9606
QY	1740	TGTCTCTATGACCGGATTCGCAAAAGCAAGTGAATTCGTCGGAATATGAGTGGCAAGATCC	1799
Dp	9607	TGTCTCTATGACCGGTTTCGCAAAAGCAAGTCTATTTGCTCGGAATATGAGCCGCAAGATCC	9666

RESULT 15	
LOCUS	AVU83986
DEFINITION	Agrobacterium vitis plasmid pTIG474 agrocinopine synthase, (5), and (1aah) gene, (1aah) pseudogene, (1pt) and (6b) genes, and (ocs) pseudogene, complete cds.
ACCESSION	U83986
VERSION	U83986.1
KEYWORDS	GI:1814315
SOURCE	Agrobacterium vitis (Rhizobium vitis)
ORGANISM	Agrobacterium vitis
REFERENCE	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
AUTHORS	1 (bases 1 to 14960)
TITLE	Otten, L. and De Ruffray, P.
JOURNAL	Sequence and functional analysis of the shoot-inducing octopine T-DNA of Agrobacterium vitis strain CG474
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 14960)
TITLE	Otten, L. and De Ruffray, P.
JOURNAL	Direct Submission
REFERENCE	Submitted (07-JAN-1997) Institute for Plant Molecular Biology, rue du General Zimmer 12, Strasbourg 67084, France
FEATURES	Location/Qualifiers
SOURCE	1..14960
	/organism="Agrobacterium vitis"
	/mol_type="genomic DNA"
	/strain="CG474"
	/db_xref="taxon:373"
	/plasmid="pTIG474"
	7244..8395
	/organism="Agrobacterium vitis"
	/mol_type="genomic DNA"
	/db_xref="taxon:373"
	/insertion_seq="IS870-like insertion element"
	11921..12767
	/organism="Agrobacterium vitis"
	/mol_type="genomic DNA"
	/db_xref="taxon:373"
	/insertion_seq="IS869"
	252..14745
	/note="shoot-inducing octopine T-DNA"
	252..278
	/note="flanking repeat; left border of the T-DNA"
	/rpt_type=direct
	complement (496..1854)
	/codon_start=1
	/transl_table=1
	/product="agrocinopine synthase"
	/protein_id="AA81866.1"
	/db_xref="GI:1814316"
	/translation="MALNVRVLYLLTMTCHNIPPEFAQVLCQRKPKVEGTWPAWTL
	LHMDPENVTPLDLVAHSVPRLLELEEDLPQVIEHRGMFNIGKIOCTASSLLA
	GLGGEENLSSEMDVCLTSDNVAIVSHDLNTRVSEKGDLPFNEIHSSIKDYVIR
	EVANGNLODKVPFETVDHIPSLETLIGKVFANPDATIFLGRNVEAHVIAAMSHREE
	YHQRVVLPTFEYFHGAGVDAVIAKQAPSDMRKSIALMPALPRELCGLARLROVT
	EPVAVDDYVLAKGAMIDSMQWRIVAAHVFSGVTRNLGGVAVDQVLAAPFSEDA
	VRVLAIVKEDDTMRAKRPHLKFAAVTRCYDFALLDGEGERFSIDIKTRARHREDA
	ERKIRWRKCTPGNSATIAWVISDRPEDEMAIWEWNQIDREVSHTLPHLVSVEI
	PG"
	2363..3046
	/gene="5"
	2363..3046
	/gene="5"
	/codon_start=1
	/transl_table=1
	/protein_id="AA81867.1"
	/db_xref="GI:1814317"

```

gene
CDS
    /translation="MWSRPINVDISNTHRCRRLKXFLRHVEIAYSFAPORLIRVA  

    OSLSIINIINDVIDPAIDDEVNRCCEVACQAGAGIPLRIILNDSIMTYCSFRRK  

    RRYAHKRLIDGSDSVIATVPYPTTEGIRKRTSRKQNNNGENTSDIENVDYIAFA  

    LPTSLQNSVSHMKIRCNLSFAPSRVDFCVELLIAVGKALFEDNGLRKKRLWGENMA  

    LTLWMKRLV"  

    complement(3569..4972)  

    /gene="iaah"  

    complement(3569..4972)  

    /gene="iaah"  

    /codon_start=1  

    /transl_table=1  

    /protein_id="AAB41868.1"  

    /db_xref="GI:1814318"  

    /translation="MWSITSLAOSLEHLRRDYSCLEIVETLIARCEAKSLMNLAT  

    DWDLGRLSKKITDRHGNAGVGLCGIPLCEKANIATGVPTSLATPALINHLPKIISRVR  

    AARLPSAGALPGASGMHGLSPGITSNNYATGAVNPNMPSLISGSSGGYAAAVASG  

    LMLGCGTGTGASVRLPALCGVGFPRPLIGRPDRILIPSPPTDTAGIYAOCVADY  

    ILLDVOISRPAPRIPIPLKGIKRLPTTYPFDLDADVAFAEETIRLLANRVTYKKA  

    EADIPHEXRXGASLPIALYEFPHALQIDDPGATYSFSDVINGISHPVANIYKKA  

    QYDQHOISDEVELARQSFRRPLQATRYNFFLYOLDILPPFAVLAKAIGDQFSVAT  

    NNGSTMTNTEKIRIVRVNDPSSNAGLPGLSLPHVLTBRLPVGMEIDGLAGSHRLLAIG  

    MALEKRIIDHLEPNVYLN"  

    5306..8720  

    /gene="iaam"  

    join(5306..7243,8396..8720)  

    /gene="iaam"  

    /note="Interrupted by a bacterial IS870-like insertion  

    element"  

    /pseudo  

    /codon_start=1  

    /transl_table=1  

    7244..7247  

    /note="flanking repeat; left border of the IS870-like  

    insertion element"  

    /rpt_type=direct  

    7301..8395  

    /gene="ORF1"  

    7301..8395  

    /gene="ORF1"  

    /codon_start=1  

    /transl_table=1  

    /protein_id="AAB41869.1"  

    /db_xref="GI:1814319"  

    /translation="MRIMCMRTGVSFTVSPTDHQRMLALTRDNAPOKHWRSSEIILL  

    SADGVGYEIMRQTSKTCVWRQERFAAEVSEGLSDKTRPSRIYALGPDVAFERVXQ  

    AARLLIRARROKHTADMMATVAGISASAVRIRWAGHGLAHRMOPFLSNDPFEVXD  

    LADVLGVYDPAVAHIVLSVDEKSOIQALDPTQPLPMKKRIGTMTHDYKRRHTTLL  

    PLAVLVLDCTVYCGMQRHROEIRFLKRIINAQPPDDKSFHVILDAINAKHKVYKA  

    WLERBERLFTHTPTSCGMNANVEGFARLSFRLLKRGVFSVVDLQAINRFLVEHNI  

    KKKPFTVADPDKIIAAVXKROHQLDSIH"  

    8395..8395  

    /note="flanking repeat; right border of the IS870-like  

    insertion element"  

    /rpt_type=direct  

    9419..10141  

    /gene="ipt"  

    9419..10141  

    /gene="ipt"  

    /codon_start=1  

    /transl_table=1  

    /protein_id="AAB41870.1"  

    /db_xref="GI:1814320"  

    /translation="MDLRLFGPTCTGKTSFAIRLAQOGLPVLSDRVQCCPOLSTGSG  

    SCRPTVEELKGTGTRLLYEDRLPLVKGIIAKQKHRLIGEVYNEAHGGLIEGSIISL  

    LRCSMOSXWSTSDPFRWHILIRHKLADERTFMAAKRANQOMLRPAVGSLITQELVYNN  

    ERRLRLPLKELIDGIRYAMLFAQONQITPDMLLQDDPMNEGELHGIABOYLIHAROE  

    QSEFPVSVAAFEGFSGPPGMC"  

    complement(10816..11439)  

    /gene="cb"  

    complement(10816..11439)  

    /gene="cb"  

    /codon_start=1  

    /transl_table=1

```

	gene	/protein_id="AAB41871.1" /db_xref="GI:1814321" /translation="MTVAAMQVRDLTLILSAGEIQSLERQARDTFGLGDI VYFQPES VISLGEFDDEIVIFSGDELVVVYLVEDIAROCALRRNPSPSSNNGI VATAPPMIMNA HGNVLQJLRDCPGLVNVYIGEPHTNOVFPLAIMSNFSVAIVINNESYGFPARGC NYTEAGAGGDVEVDDEGAAATTESQIGNLVIPVALGSCQLSA"					
	CDS	/gene="Ocs" complement(13214..14290) /pseudo /codon_start=1 /transl_table=11 14719..14745 /note="flanking repeat; right border of the T-DNA" /rpt_type=direct					
	repeat_region						
BASE COUNT	3854 a	3510 c	3632 g	3958 t	6 others		
ORIGIN							
Query Match	81.3%; Score 1464.4;	DB 1;	Length 14960;				
Best Local Similarity	89.0%;	Pred. No. 0;					
Matches 1604;	Conservative 0;	Mismatches 196;	Indels 2;	Gaps 2;			
Qy	1	ATGTCACTTCACCTCTCTCTTGAATAACCAGTGCATCATCTCCCAACCAAAATGTGGAT	60				
Db	5306	ATGTCACTTCACGCTCTCTTGAATATCCAAGCATATTTCTTACCAAAATGTGGAT	5365				
Qy	61	CTGCACAATGTGTCTAAGCGGATGAATTGGACCGCAGAGGTTTTCCGATGCCCTTCTTAGAA	120				
Db	5366	CTGATTAATGTGTCTAAGCTCTGAAGTGTGGCCGCGGGTGTTGGCGATGCTTCTTCAGAG	5423				
Qy	121	CGAGAAGCTTCTAGAGGGAAGAGATTACTCAAAGCTCCACCGATGCAGCGCTGGGTTA	180				
Db	5426	CGAGAATCTTCTAGAGGCAATGAGAAATTACTCAAATTTCCGGCGATGCGAGCGCTGGGTTA	5485				
Qy	181	GCTTGCAAAAGGCTGGCCGATGTCCTTCCCAGATCTTCACTGTGTGAAAGTATGCA	240				
Db	5486	GCTTGCAAAAGGCTGGCCGATGTCCTTCCCAGATCTTCAAGCGGTGAGATGTATGCA	5545				
Qy	241	GTTCCTCCGCTTATCTATATTTGGGCAAGAAATCTGGGGCCGATATCTGAATTCGAAA	300				
Db	5546	GTCTCTCCCCGTTTATCTATATTTGCTGGGAGAGAAATTTGGGCCCGATATCTGAATTCGAAA	5605				
Qy	301	CCTTGGCGCGGCGCAACAGTAGAGTGTCTCGTCCATCGACTTGAACACCATTTGCATG	360				
Db	5606	CAATGGCGCGGCAAGTAGAGTGTGTCTGTCATCGACTTGAACCATTTTGTATG	5665				
Qy	361	GATTTTCTCCGAAGCACAACTTAATCAAGCCCTGTTTTTGTGAGCGGTAAAAAGTATGCA	420				
Db	5666	GATTTTCTCCGAAGCACAACTTCTCAAAATCTGAATTTGTGAGCGGCAAAAGATATGGA	5725				
Qy	421	CCGATTAATCTTAATGTATCTTGGAGCAATTTCAATCTCTAAGATGTGCGGCTTTCGAAC	480				
Db	5726	CCGAGCACTTAATGTATCTTGGAGCAATTTCTGAAGTGTGCGGCTTTCGGAAC	5785				
Qy	481	CTGCGCAATGTCCGCTGTATCGAGAAATGGACGATGAATATGCGGTATCCGATTAAC	540				
Db	5786	CTGCAAAATGTCCGCGGTATCGAGAACGCGATGAAGACGCTTAATGGGTTTGCATGACC	5845				
Qy	541	CTTGAAGGGGCGGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	600				
Db	5846	CTTGAAGGGGCGGTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	5905				
Qy	601	TGCGGAGGTTCTTTCCAAACAATGCAATTTGCTGACGACGTGACAGACGTTTTCGACCA	660				
Db	5906	TGCGGAGGTTCTTTCCAAACAATGCAATTTGCTGACGACGTGACAGACGTTTTCGACCA	5965				
Qy	661	TGTTCCGATAGTGAACGATGCGCTTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG	720				
Db	5966	TGTTCCGATAGTGAACGATGCGCTTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG	6025				
Qy	721	GTCATTTGGCGCTGGCATTTCCGACTCTGTGTGCAACGAACTGTTTCAATGCTGGGCTA	780				

Search completed: November 24, 2003, 01:37:14  
Job time : 6650.53 secs

Db	6026	GTCAATTGGGGCTGGGCATTTTCCGGGGCTCGTGGTGGCAAGCAACTGCTTCAATGCTGGTGTA	6085
Qy	781	GAGCAGTGTTCACATATATATGAAAGCAAGTGATCGTGTGGAGGCAAGCTTTGGTCAATGCT	840
Db	6086	GACGATGTTACAAATATATGAAAGCAGTGATCGGGTTGGAGGCAAGCTTTGGTCAATGCT	6145
Qy	841	TTTCAGGCAAGCTTCTTAGTGTGTTGGCCGAAATGGGGGCGATGCGATTTTCTCCCTGCTGCA	900
Db	6146	TTTCAGGAAAGCTCGAGAGGCTCAAGCCGAAATGGAGGAGATGCAATTTTCTCTGCTGCA	6205
Qy	901	TTTCGCTGTGTTTTTCTTCTCTCGAGCGTTATCGGCCTGTCTTTCGATGAGGCGCTTCCCAAT	960
Db	6206	TCGTCGTTGTTCTTCTCTCTCGAGCGGTACGSGTGTCTGCTTCGATGAGGCCATTTCCCAAT	6265
Qy	961	CCGGGCAACGTGCACTTACTTGGTCTTACCAAGCGCTCCAAATCATGTGAAAGCCGGG	1020
Db	6266	CCGGGCAACGTGCACTTACTTGGTCTTACCAAGCGGTCCAAATCATGTGAAAGCCGGG	6325
Qy	1021	CAGCTGCCACCGAAGCTGTTCATTCGCGTTTACAAAGCTTTGGCGGTCTTGAAGAC	1080
Db	6326	CAGCGCGCACGGAAGTGTTCATTCGCGTTTACAAAGCGGTGCGATGCTTCTTAAAGAC	6385
Qy	1081	GGTTTTTCAGCGGAGATATGTGTGGCTTCGCGCTGTCGCTATTACTCAGGCGTTGAA	1140
Db	6386	GGTTTTCTTGGGAGGAGATATGTGTGGCTTCGCGCTGTCGCTATTACTCAGGCGCTGAA	6445
Qy	1141	TCAGGACACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTTGGGAGG	1200
Db	6446	TTAGGGGACATTTAAGCGGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTTGGGAGG	6505
Qy	1201	GAGTCTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGCAACATCTCTCTGTGTGT-AA	1259
Db	6506	GAGTCGTTCTCTTCAAGCATAGAGAGATCTTTCTGGGCTGCAATCTCTCGGTGTGTGAA	6565
Qy	1260	ACATGGAATTTTCCCATGATGTTGGGACCTATTCAAGCTTAATGGGAATAGATCTGGCGGG	1319
Db	6566	ACATGGAATTTTCCCATGATGTTGGGACCTATTCAAGCTTAATGGGAATAGATCTGTGC-GG	6624
Qy	1320	GTTTGTCTCAGTTTTTGAAGCGGGTTTATTAAGATCTCTCGCTTGTCATCAACGATA	1379
Db	6625	GTTTGTCTCAGTTTTTGAAGCGGGTTTATTAAGATCTCTCGCTTGTCATTAACGATA	6684
Qy	1380	TGAAGAAATCAGCGGATGTGCTCTGAAGGAATCTCAGAACTTTCACGTCGGATTCGATC	1439
Db	6685	TGAAGAAATCAGCGGATGTGCTCTGAAGGAATCTCAGAACTTTCACGTCGGATTCGATC	6744
Qy	1440	TGAAGTGTTAACGGTGTGTCTGTGAGCGAGCGCATATGCCATGTTCAAGTCAGGGCGAT	1499
Db	6745	TGAAGTGTTAAGGTGTGTCTGTGAGCGAGCGCATATGCCATGTTCAAGTCAGGGCGAT	6804
Qy	1500	TCAGAGGAAAGACAAAAATTAAGATTAAGCTTAAGACCGGATATCTGAATCTTTATGA	1559
Db	6805	TGAAGAGAAAGACAAAAATTAAGATTAAGCTTAAGACCGGATATCTGAATCTTTATGA	6864
Qy	1560	TAAAGTGTGTGCATCTGGAATTCGCAATATCCAACTCAGGCAATTCCTGCAATCGGA	1619
Db	6865	TAAAGTGTGTGCATCTGGAATTCGCAATATCCAACTCAGGCAATTCCTGCAATTCGA	6924
Qy	1620	TACCAATATTTTTTCAGGACAGGAGAACCAAGCGGTTATTAACGCCATATGACAGATC	1679
Db	6925	TACCAATATTTTTTCAGGACAGGAGAACCAAGCGGTTATTAACGCCATATGACAGGCTC	6984
Qy	1680	GTCAAAATCTTCCGATGATCTGAAAGCAAAATTCGTGTTAAGCAATATCTCCGCTTGG	1739
Db	6985	TTCAAAAATCTTCTTGTGCTGAAGCAAAATTTTGGATCAACGATATGCTCCGCTCTG	7044
Qy	1740	TGTCCTCATGACGGGATCGCAAAAGCAGTGTATTCCTGGAATATGAGTGCAGAGATCC	1799
Db	7045	TGTCCTCATGACGGGATCGCAAAAGCAGTGTATTCCTGGAATATGAGTGCAGAGATCC	7104
Qy	1800	GA 1801	
Db	7105	GA 7106	

PT Gall resistant plants generated by transforming cells with an

TL-DNA region of A  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Nucleic acid encod  
Aspergillus oryzae  
Human secreted pro  
DNA encoding novel  
DNA encoding novel  
Propionibacterium  
Mycobacterium tub  
Mycobacterium tub  
Micrococcus luteus  
Mutine secreted pr  
Human secreted pro  
Human secreted pro  
Human secreted pro  
Human polyamine ox  
Human polyamine ox  
Human polyamine ox  
Human polyamine ox  
P acnes linoleate  
P acnes linoleate  
P acnes linoleate  
Saccharomyces cere  
Saccharomyces cere  
Cell cycle protein  
Cell cycle protein  
Bifidobacterium jo  
Tobacco plant resi  
Human transporter  
Hydrophobic domain  
Human polyamine ox  
Human polyamine ox  
Human polynucleoti  
Human polynucleoti  
Human oxidoreducta

PT untranslatable nucleic acid homologous to a gall disease-causing gene

XX Claim 2; Page 45; 49pp; English.

CC The patent discloses a method for producing plants resistant to gall  
CC disease, induced by Agrobacterium infection, by transformation with a  
CC bacterial resistance (BR) construct capable of eliciting co-suppression.  
CC The construct encodes an untranslatable RNA molecule, that is highly  
CC homologous to the tumour or gall disease-causing genes *Jaam*, *Jaah* and  
CC *1st* of *Agrobacterium tumefaciens*. These oncogenes are modified by  
CC introduction of premature termination codons or frameshift mutations, to  
CC inhibit their expression. This method is used to reduce susceptibility  
CC of plants to gall diseases. The present DNA sequence is the *Agrobacterium*  
CC *tumefaciens* T-DNA (transferred DNA) gene *Jaam*, required for auxin  
CC production. It encodes an enzyme that converts tryptophan into indole  
CC acetylamine, essential for auxin synthesis. The overproduction of these  
CC plant growth hormones results in gall formation.

XX Sequence 2268 BP; 543 A; 520 C; 606 G; 599 T; 0 other;

Query Match 100.0%; Score 1801; DB 21; Length 2268;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCACTTCACTCTCTTGTATACCAAGCGCATCATCTCCCAACCAAAATGTGAT 60  
DB 1 ATGTCACTTCACTCTCTTGTATACCAAGCGCATCATCTCCCAACCAAAATGTGAT 60  
QY 61 CTGCAATGTGATGATTAAGCGCGATGATTTGACCGAGGTTTCCGATGCTTTAGAA 120  
DB 61 CTGCAATGTGATGATTAAGCGCGATGATTTGACCGAGGTTTCCGATGCTTTAGAA 120  
QY 121 CGAAGACCTTCTAGGGAAGAGGATTAATCTCAAGCTTCCAGTGCAGCGCTGGGTTA 180  
DB 121 CGAAGACCTTCTAGGGAAGAGGATTAATCTCAAGCTTCCAGTGCAGCGCTGGGTTA 180  
QY 181 GCTTGCAAAAGCGCGCGATGATGCTTCCCGAGATCTCAAGCTTCCAGTGCAGCGCTGGGTTA 240  
DB 181 GCTTGCAAAAGCGCGCGATGATGCTTCCCGAGATCTCAAGCTTCCAGTGCAGCGCTGGGTTA 240  
QY 241 GTTCTCTCCGCTTATCTATATTTGCAAAAGATTTCTGGGCGGATCTTGAATCGAAA 300  
DB 241 GTTCTCTCCGCTTATCTATATTTGCAAAAGATTTCTGGGCGGATCTTGAATCGAAA 300  
QY 301 CTTTGGCGCGGCAACAGTGAAGTGTCTGCTTCCATTCGACTTGAACCATTTGCATG 360  
DB 301 CTTTGGCGCGGCAACAGTGAAGTGTCTGCTTCCATTCGACTTGAACCATTTGCATG 360  
QY 361 GATTTCCTCCGAAGCAACTAATCCAAAGCCCTGTTTTGTGAGCGGTTAAAGATGCA 420  
DB 361 GATTTCCTCCGAAGCAACTAATCCAAAGCCCTGTTTTGTGAGCGGTTAAAGATGCA 420  
QY 421 CCGATTGATCTTACTGATCTTGTGCGCATTTCAATCTCTAAGCTGCGGCTTTCGAAC 480  
DB 421 CCGATTGATCTTACTGATCTTGTGCGCATTTCAATCTCTAAGCTGCGGCTTTCGAAC 480  
QY 481 CTGCGAATGCGCGTGTACGAGAAATGSCACGATGAATGCGTTACCGGGTTTACATAAC 540  
DB 481 CTGCGAATGCGCGTGTACGAGAAATGSCACGATGAATGCGTTACCGGGTTTACATAAC 540  
QY 541 CTTGAAGAGGCGCGCGCATTTGACATGATGATGATGATGATGATGATGATGATGAT 600  
DB 541 CTTGAAGAGGCGCGCGCATTTGACATGATGATGATGATGATGATGATGATGATGAT 600  
QY 601 TCGGAGAGTTCCTTCCAAACATGATGATGATGATGATGATGATGATGATGATGAT 660  
DB 601 TCGGAGAGTTCCTTCCAAACATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 661 TGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
DB 661 TGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

QY 721 GTGATTGGCGCGCGCATTTCCGAGCTGTGTGCGCAACAGCACTGCTCATGCTGGGTA 780  
DB 721 GTGATTGGCGCGCGCATTTCCGAGCTGTGTGCGCAACAGCACTGCTCATGCTGGGTA 780  
QY 781 GACGATGTTACAAATATATGAAAGCAAGTATGATGATGATGATGATGATGATGATGAT 840  
DB 781 GACGATGTTACAAATATATGAAAGCAAGTATGATGATGATGATGATGATGATGATGAT 840  
QY 841 TTCAAGGACGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 841 TTCAAGGACGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 901 TTGCGCTGTTTCTTCTCCGAGCGCTTACGCGCTGCTGATGATGATGATGATGATGAT 960  
DB 901 TTGCGCTGTTTCTTCTCCGAGCGCTTACGCGCTGCTGATGATGATGATGATGATGAT 960  
QY 961 CCCGACAGTGCACACTTACTTGTCTACCAAGCGCTTCAATACATGATGAAAGCCGGG 1020  
DB 961 CCCGACAGTGCACACTTACTTGTCTACCAAGCGCTTCAATACATGATGAAAGCCGGG 1020  
QY 1021 CAGCTGCCACCGAAGCTTCTTCCATCCGCTTACCAAGGTTGGCGCTTCTTGAAGAC 1080  
DB 1021 CAGCTGCCACCGAAGCTTCTTCCATCCGCTTACCAAGGTTGGCGCTTCTTGAAGAC 1080  
QY 1081 GGTTCATGAGGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
DB 1081 GGTTCATGAGGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
QY 1141 TCAAGACATTAAGTGGGCTCATGATCTCTGCAAAATTTGGCTGAACCGTTTGGGAGG 1200  
DB 1141 TCAAGACATTAAGTGGGCTCATGATCTCTGCAAAATTTGGCTGAACCGTTTGGGAGG 1200  
QY 1201 GAGTCTTCTCTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
DB 1201 GAGTCTTCTCTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
QY 1261 CATGAGATTTCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
DB 1261 CATGAGATTTCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
QY 1321 TTTGTCACATTTTGAAGCGGATTTATGAGATCTCTCGCTTGTGATCAACGATAT 1380  
DB 1321 TTTGTCACATTTTGAAGCGGATTTATGAGATCTCTCGCTTGTGATCAACGATAT 1380  
QY 1381 GAAAGAAATACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1381 GAAAGAAATACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
QY 1441 GAAAGGTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
DB 1441 GAAAGGTTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
QY 1501 CAGAGGAAAGCAAAATTAAGCTTAAGGCTTAAGGCGGATATCTGAACCTTTATGAT 1560  
DB 1501 CAGAGGAAAGCAAAATTAAGCTTAAGGCTTAAGGCGGATATCTGAACCTTTATGAT 1560  
QY 1561 AAGGTGTGTCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
DB 1561 AAGGTGTGTCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
QY 1621 ACCAATTTTTCAGGACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATTCG 1680  
DB 1621 ACCAATTTTTCAGGACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATTCG 1680  
QY 1681 TCAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
DB 1681 TCAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
QY 1741 GTCTCATGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
DB 1741 GTCTCATGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
QY 1801 A 1801

Db 1801 A 1801

## RESULT 2

AAD00625 standard; DNA; 1807 BP.

AAD00625;

29-AUG-2000 (first entry)

Agrobacterium tumefaciens T-DNA mutant iaam gene.

iaam; iaah; ipt; T-DNA; transferred DNA; gall-disease resistance; BR; bacterial resistance construct; untranslatable RNA; tumour; oncogene; auxin; plant growth hormone; mutant; ds.

Agrobacterium tumefaciens.

WO200026346-A1.

11-MAY-2000.

04-NOV-1999; 99WO-US26100.

05-NOV-1998; 98US-0107185.

(UYOR-) UNIV OREGON STATE.

Ream W, Mok MC, Lee H;

WPI; 2000-365599/31.

Gall resistant plants generated by transforming cells with an untranslatable nucleic acid homologous to a gall disease-causing gene

Example C; Page 43-44; 4ppp; English.

The patent discloses a method for producing plants resistant to gall disease, induced by Agrobacterium infection, by transformation with a bacterial resistance (BR) construct capable of eliciting co-suppression. The construct encodes an untranslatable RNA molecule, that is highly homologous to the tumour or gall disease-causing genes iaam, iaah and ipt of Agrobacterium tumefaciens. These oncogenes are modified by introduction of premature termination codons or frameshift mutations, to inhibit their expression. This method is used to reduce susceptibility of plants to gall diseases. The present DNA sequence is the Agrobacterium tumefaciens T-DNA (transferred DNA) mutant iaam gene. This oncogene was modified by introduction of a stop sequence (TGA) at the third codon and deletion of two bases following the third codon. This sequence was used for the construction of double stranded RNA and untranslatable dsRNA vectors, that reduce gall formation in transformed plant cells.

Sequence 1807 BP; 428 A; 418 C; 486 G; 475 T; 0 other;

Query Match 99.3%; Score 1788; DB 21; Length 1807;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12 ACCCTCTCTGATACAGAGCGGATCATCTCCCAACCAAAATGGTATCTGCACAAATGT 71  
18 ACCCTCTCTGATACAGAGCGGATCATCTCCCAACCAAAATGGTATCTGCACAAATGT 77  
72 CGATTAAGCGGATGAATTTGACCGCAGGTTTCGATGCTTCTTAAGACGAGAGCTTC 131  
78 CGATTAAGCGGATGAATTTGACCGCAGGTTTCGATGCTTCTTAAGACGAGAGCTTC 137  
132 TAGGGGAGAGAGATTAATCTCAAGCTCCACGAGTGACGCGTGGTTAGCTTGCAAAAG 191  
138 TAGGGGAGAGAGATTAATCTCAAGCTCCACGAGTGACGCGTGGTTAGCTTGCAAAAG 197

QY 192 GCTGGCCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGAGTAGACGTTCTCCGC 251  
DB 198 GCTGGCCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGAGTAGACGTTCTCCGC 257  
QY 252 TTATATCTAATATTGGCAAGAAATTTCTGGGCGGATACCTTGAATGAAACCTTGGCGCG 311  
DB 258 TTATATCTAATATTGGCAAGAAATTTCTGGGCGGATACCTTGAATGAAACCTTGGCGCG 317  
QY 312 GCGAAGAGTAGAGTCTCGTTGCGATCGACCTTGGACACCACTTGGATGATTTCTCGCA 371  
DB 318 GCGAAGAGTAGAGTCTCGTTGCGATCGACCTTGGACACCACTTGGATGATTTCTCGCA 377  
QY 372 AGCACAATTAATCAAGCCCTGTTTGTCTGAGCGGTAAGAAAGATGTSCACCGATTGATCT 431  
DB 378 AGCACAATTAATCAAGCCCTGTTTGTCTGAGCGGTAAGAAAGATGTSCACCGATTGATCT 437  
QY 432 TAGTCATTTGCTGGCCATTTCAATCTCTAAGACTGCCGGCTTTGCAACCTGCGCAATGCC 491  
DB 438 TAGTCATTTGCTGGCCATTTCAATCTCTAAGACTGCCGGCTTTGCAACCTGCGCAATGCC 497  
QY 492 GCTGTAGAGAAATGGCAACGATGAAATGCCGTTACCGGGTTACCATACCTTGAAGGGCG 551  
DB 498 GCTGTAGAGAAATGGCAACGATGAAATGCCGTTACCGGGTTACCATACCTTGAAGGGCG 557  
QY 552 CGTGCCATTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 611  
DB 558 CGTGCCATTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 617  
QY 612 CTTTCCAAATCGACTTGTCTACAGCTGACAGCCGTTTCTTGTGACCAATGTTCCGATAG 671  
DB 618 CTTTCCAAATCGACTTGTCTACAGCTGACAGCCGTTTCTTGTGACCAATGTTCCGATAG 677  
QY 672 TGGACGATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCGCTCATTTGCGCG 731  
DB 678 TGGACGATGCGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCGCTCATTTGCGCG 737  
QY 732 TGGCATTTCCGACCTGCTGTGGCAACGAACTGCTTCACTGCGGGTAGACATGTTAC 791  
DB 738 TGGCATTTCCGACCTGCTGTGGCAACGAACTGCTTCACTGCGGGTAGACATGTTAC 797  
QY 792 AATATATGAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 851  
DB 798 AATATATGAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 857  
QY 852 TCTAGTGTGTGCGCGAAGATGGGGCGATGCGATTTCTCTCTGCTGATTTGCTTGT 911  
DB 858 TCTAGTGTGTGCGCGAAGATGGGGCGATGCGATTTCTCTCTGCTGATTTGCTTGT 917  
QY 912 TTTCTTCTGAGCGTTAGCGGCTGTCTTGAAGAGCGGTTCCCAATCCCGGACAGT 971  
DB 918 TTTCTTCTGAGCGTTAGCGGCTGTCTTGAAGAGCGGTTCCCAATCCCGGACAGT 977  
QY 972 CGACACTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1031  
DB 978 CGACACTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1037  
QY 1032 GAAGCTGTTCAATGCGGTTTACAAGGTTGCGTGGCTTCTTGAAGAGCGGTTTATGA 1091  
DB 1038 GAAGCTGTTCAATGCGGTTTACAAGGTTGCGTGGCTTCTTGAAGAGCGGTTTATGA 1097  
QY 1092 GCGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1151  
DB 1098 GCGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1157  
QY 1152 TAGTGGGCTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1211  
DB 1158 TAGTGGGCTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1217  
QY 1212 TTCAAGGATAGAGAGATCTTCTGCGCAACATCTCTGCTGCTGATGAATGATGATGATGAT 1271  
DB 1218 TTCAAGGATAGAGAGATCTTCTGCGCAACATCTCTGCTGCTGATGAATGATGATGATGAT 1277  
QY 1272 CTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1331

```
Db 1278 CCTCATATTGGGACCTATTCAAGCTAAATGGGAATAGATCTGGCGGGTTTGTCACGT 1337
Qy 1332 TTTTGAAGCGGGTTTATGAGATCTCCGCTTGTCATCAAGGATATGAAGAAATCA 1391
Db 1338 TTTTGAAGCGGGTTTATGAGATCTCCGCTTGTCATCAAGGATATGAAGAAATCA 1397
Qy 1392 GCGGATGTGCGCTGAAGGAATCTCAGAACTTCCAGTCGTGCGATCCGATCTGAAGTGTTAA 1451
Db 1398 GCGGATGTGCGCTGAAGGAATCTCAGAACTTCCAGTCGTGCGATCTGAAGTGTTAA 1457
Qy 1452 CGGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAGATCAGGCGCATTCAGAGAAAA 1511
Db 1458 CGGTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAGATCAGGCGCATTCAGAGAAAA 1517
Qy 1512 GACAAAATAAAGATTAAGCTTAAAGACCGGAGATCTGAACCTTTATGATAAAGTGTTGT 1571
Db 1518 GACAAAATAAAGATTAAGCTTAAAGACCGGAGATCTGAACCTTTATGATAAAGTGTTGT 1577
Qy 1572 CACATCTGGAAGCTGCAATATCCAACTCAGGCATTTGCTGACATGCGATACCAATATTTT 1631
Db 1578 CACATCTGGAAGCTGCAATATCCAACTCAGGCATTTGCTGACATGCGATACCAATATTTT 1637
Qy 1632 TCAGGACACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATCGTCAAACTCTT 1691
Db 1638 TCAGGACACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATCGTCAAACTCTT 1697
Qy 1692 CCGTATGACTGAAGCAAAATTTCTGTTAGACCATATCTCCCGTCTGTGTCTCATGA 1751
Db 1698 CCGTATGACTGAAGCAAAATTTCTGTTAGACCATATCTCCCGTCTGTGTCTCATGA 1757
Qy 1752 CGGGATGCAAAAGCAAGTATGCTCGAGCTATGAGTCCGAGATCC 1799
Db 1758 CGGGATGCAAAAGCAAGTATGCTCGAGCTATGAGTCCGAGATCC 1805

RESULT 3
AAD00626
ID AAD00626 standard; DNA; 2544 BP.
AC AAD00626;
XX
DT 29-AUG-2000 (first entry)
XX
DE Agrobacterium tumefaciens ipt/iaam untranslatable BR construct.
XX
KM iaam; iaah; ipt; T-DNA; transferred DNA; gall-disease resistance; BR;
KM bacterial resistance construct; untranslatable RNA; tumour; oncogene;
KM plant growth hormone; ds.
XX
OS Agrobacterium tumefaciens.
XX
FH Key 1..309 Location/Qualifiers
FT misc_feature /*tag= a
FT /*note= "Corresponds to Agrobacterium mutant ipt gene"
FT /*tag= b
FT /*note= "Corresponds to Agrobacterium mutant iaam gene"
XX
PN WO200026346-A1.
XX
PD 11-MAY-2000.
XX
PF 04-NOV-1999; 99WO-US26100.
XX
PR 05-NOV-1998; 98US-0107185.
XX
PA (UYOR-) UNIV OREGON STATE.
XX
PI Ream W, Mok MC, Lee H;
XX
DB WPI, 2000-365599/31.
```

```
XX Gall resistant plants generated by transforming cells with an
PT untranslatable nucleic acid homologous to a gall disease-causing gene
PT
PS Example c; Page 44-45; 49pp; English.
XX
CC The patent discloses a method for producing plants resistant to gall
CC disease, induced by Agrobacterium infection, by transformation with a
CC bacterial resistance (BR) construct capable of eliciting co-suppression.
CC The construct encodes an untranslatable RNA molecule, that is highly
CC homologous to the tumour or gall disease-causing genes iaam, iaah and
CC ipt of Agrobacterium tumefaciens. These oncogenes are modified by
CC introduction of premature termination codons or frameshift mutations, to
CC inhibit their expression. This method is used to reduce susceptibility
CC of plants to gall diseases. The present DNA sequence is the Agrobacterium
CC tumefaciens bacterial resistance construct, that encodes untranslatable
CC RNA forms of both the mutant ipt and iaam genes. This construct can
CC suppress both shooty and necrotic gall formation in transgenic plants.
XX
SQ Sequence 2544 BP; 625 A; 592 C; 674 G; 653 T; 0 other;
XX
Query Match 99.3%; Score 1788; DB 21; Length 2544;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 ACCTTCCTTGATTAACCAAGTGCATATCTCCCAACCAAAATGATGATCTGCAATGCT 71
Db 318 ACCTTCCTTGATTAACCAAGTGCATATCTCCCAACCAAAATGATGATCTGCAATGCT 377
Qy 72 CGATTAAGGCGGATGAATTTGGAACCGCAGGTTTCCGATGCTCTTGAACGAGAGCTTC 131
Db 378 CGATTAAGGCGGATGAATTTGGAACCGCAGGTTTCCGATGCTCTTGAACGAGAGCTTC 437
Qy 132 TAGGGGAAGGAGGATTACTCAAAAGCTCCACCGAGTCAAGCGCTGGTTAGCTTGAAG 191
Db 438 TAGGGGAAGGAGGATTACTCAAAAGCTCCACCGAGTCAAGCGCTGGTTAGCTTGAAG 497
Qy 192 GCTGGCGGATGTCGCTTCCCGAGATCTCAGTGTGTGAAAGTACATTTCTCCGC 251
Db 498 GCTGGCGGATGTCGCTTCCCGAGATCTCAGTGTGTGAAAGTACATTTCTCCGC 557
Qy 252 TTATATCTATATTTGGCAAGAAATTTCTGGGGCGGATCTTGAATCGAAACCTTGGGCG 311
Db 558 TTATATCTATATTTGGCAAGAAATTTCTGGGGCGGATCTTGAATCGAAACCTTGGGCG 617
Qy 312 GGCACAGTAGTGTCTCGTCCCATCGACTTGAACACCATTTGATGATTTCTCGA 371
Db 618 GGCACAGTAGTGTCTCGTCCCATCGACTTGAACACCATTTGATGATTTCTCGA 677
Qy 372 AGCACAATTAATCAAGCCCTGTTTTTGTGAGCGGTAAAGATGTGCACCGATTGATCT 431
Db 678 AGCACAATTAATCAAGCCCTGTTTTTGTGAGCGGTAAAGATGTGCACCGATTGATCT 737
Qy 432 TAGTCATTTGCGCCATTTCAATCTCTAAGATGTCGCGGCTTTCGAACCTGCAATGCC 491
Db 738 TAGTCATTTGCGCCATTTCAATCTCTAAGATGTCGCGGCTTTCGAACCTGCAATGCC 797
Qy 492 GCTGTACGAGAAATGGAAGTAATGCGTTTACCGGGTTTACATTAACCTTGAAGGGGC 551
Db 798 GCTGTACGAGAAATGGAAGTAATGCGTTTACCGGGTTTACATTAACCTTGAAGGGGC 857
Qy 552 CGTGCCATTTGACATGTAGCTTAATGTGGAACCTGATGCTGAAGGTTCCGCAAGTTTC 611
Db 858 CGTGCCATTTGACATGTAGCTTAATGTGGAACCTGATGCTGAAGGTTCCGCAAGTTTC 917
Qy 612 CTTTCAACATGACCTGTCTACGACGTCAGACCGTCTTTTGAACCAATGTTCCGATAG 671
Db 918 CTTTCAACATGACCTGTCTACGACGTCAGACCGTCTTTTGAACCAATGTTCCGATAG 977
Qy 672 TGAGCGATCGGCTTTCTTCGGAGAGATGTTCTTAAGCCGAAAGTGCGGTCATTGGCGC 731
Db 978 TGAGCGATCGGCTTTCTTCGGAGAGATGTTCTTAAGCCGAAAGTGCGGTCATTGGCGC 1037
```

QY 732 TGGCATTTCCGACTCGTGGTGCAAGCAACTGCTTCATGCTGGGTAGACGATGTTAC 791  
 DB 1038 TGGCATTTCCGACTCGTGGTGCAAGCAACTGCTTCATGCTGGGTAGACGATGTTAC 1097  
 QY 792 AATATATGAAGCAAGTGCATGCTGTGGAGGCAAGCTTTGCTCAATGCTTTCAAGGACGC 851  
 DB 1098 AATATATGAAGCAAGTGCATGCTGTGGAGGCAAGCTTTGCTCAATGCTTTCAAGGACGC 1157  
 QY 852 TCTTACGTGCTGGCCGCAATTTGGGGGAGATGCGATTTTCTCTGCTGCTGATTTCTGCTTTGTT 911  
 DB 1158 TCTTACGTGCTGGCCGCAATTTGGGGGAGATGCGATTTTCTCTGCTGCTGATTTCTGCTTTGTT 1217  
 QY 912 TTTCTTCTCGACGCTTACGGGCTGTGCTTGCATGAGGCGCTTCCCAAAATCCGGGACAGT 971  
 DB 1218 TTTCTTCTCGACGCTTACGGGCTGTGCTTGCATGAGGCGCTTCCCAAAATCCGGGACAGT 1277  
 QY 972 CGACACTTACTTGTGCTTACCAAGGCGTCAATAATGTTGAAAAGCCGGGACGCTGCCACC 1031  
 DB 1278 CGACACTTACTTGTGCTTACCAAGGCGTCAATAATGTTGAAAAGCCGGGACGCTGCCACC 1337  
 QY 1032 GAAGCTTCCATCCGCTTTACACAGGTTGGCGTGGCTTTTGAAGGACGGTTTTCATGA 1091  
 DB 1338 GAAGCTTCCATCCGCTTTACACAGGTTGGCGTGGCTTTTGAAGGACGGTTTTCATGA 1397  
 QY 1092 GCGAGATATTTGTTGGCTTCCGCTGTGCTATTAATCAGGCGCTTGAATCAGACACAT 1151  
 DB 1398 GCGAGATATTTGTTGGCTTCCGCTGTGCTATTAATCAGGCGCTTGAATCAGACACAT 1457  
 QY 1152 TAGTGGGCTCATGACTCTCGGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1211  
 DB 1458 TAGTGGGCTCATGACTCTCGGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1517  
 QY 1212 TTACAGGATTAAGAGATCTTTCTGGGACACATCTCTCTGCTGCTGAACATGAGTTT 1271  
 DB 1518 TTACAGGATTAAGAGATCTTTCTGGGACACATCTCTCTGCTGCTGAACATGAGTTT 1577  
 QY 1272 CCTCATGATTTGGGACCTATTCAGCTAATGGAAATGAGATCTGGGGGTTTGGTCAAGT 1331  
 DB 1578 CCTCATGATTTGGGACCTATTCAGCTAATGGAAATGAGATCTGGGGGTTTGGTCAAGT 1637  
 QY 1332 TTTTGAAGGGGTTTATTAAGATCTCTCGCTTGTGATCAACGGAATATGAAGAAATCA 1391  
 DB 1638 TTTTGAAGGGGTTTATTAAGATCTCTCGCTTGTGATCAACGGAATATGAAGAAATCA 1697  
 QY 1392 GCGGATGTGCTCTGAAGGAATCTCAAGACTTCCAGCTCGATGCAATCTGAAGTGTAA 1451  
 DB 1698 GCGGATGTGCTCTGAAGGAATCTCAAGACTTCCAGCTCGATGCAATCTGAAGTGTAA 1757  
 QY 1452 CGGTGTGCTGTGAGCGGCGCATATGCCATGTTCAAGTCAGGGCGATTCAGAAAGAAA 1511  
 DB 1758 CGGTGTGCTGTGAGCGGCGCATATGCCATGTTCAAGTCAGGGCGATTCAGAAAGAAA 1817  
 QY 1512 GACAAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGATTAAGTGTGGT 1571  
 DB 1818 GACAAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATTTATGATTAAGTGTGGT 1877  
 QY 1572 CACATCTGGAATCTGCAAAATATCCAACTCAGGCAATTCCTGACATGCGATATCAATATTTT 1631  
 DB 1878 CACATCTGGAATCTGCAAAATATCCAACTCAGGCAATTCCTGACATGCGATATCAATATTTT 1937  
 QY 1632 TCAGGACACAGTGAACCAAGCGGTTGATTAACGCCATATGACAGGATTCGCAAAACTCTT 1691  
 DB 1938 TCAGGACACAGTGAACCAAGCGGTTGATTAACGCCATATGACAGGATTCGCAAAACTCTT 1997  
 QY 1692 CCTGATGACTGAACGAAAAATTTCTGTTAGACATATCTCCGCTTGTGCTCTCATGA 1751  
 DB 1998 CCTGATGACTGAACGAAAAATTTCTGTTAGACATATCTCCGCTTGTGCTCTCATGA 2057  
 QY 1752 CCGGATCGCAAAAGCAAGTGTATTCCTGGAATATGAGTGCAGAGATCC 1799  
 DB 2058 CCGGATCGCAAAAGCAAGTGTATTCCTGGAATATGAGTGCAGAGATCC 2105

RESULT 4  
 AAN50182  
 ID AAN50182 standard; DNA; 24596 BP.  
 XX  
 AC AAN50182;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 17-OCT-1991 (first entry)  
 XX  
 DE Complete nucleotide sequence of the T-DNA region of the octopine Ti  
 DE plasmid pTi15955.  
 XX  
 KM Plant vector; transformation-inducing principle (TIP) gene;  
 KM octopine Ti plasmid; ss.  
 OS Agrobacterium tumefaciens ATCC 15955.  
 XX  
 PN EP14538-A.  
 XX  
 PD 19-JUN-1985.  
 XX  
 PF 16-NOV-1984; 84EP-0307969.  
 XX  
 PR 18-NOV-1983; 83US-0553786.  
 PA (AGRK ) AGRIGENETICS RES ASSOC LTD.  
 PA (LUBR ) LUBRIZOL GENETICS INC.  
 XX  
 PI Barker RF, Kemp JD;  
 XX  
 DR WPI; 1985-148223/25.  
 XX  
 PT New DNA vectors contg. T-DNA sequence of octopine Ti plasmid -  
 PT for expression in plant cells to confer desirable properties to  
 PT plants and their cells  
 XX  
 PS Claim 28; Fig 1; 87pp: English.  
 XX  
 CC The inventors claim a vector contg. a transformation-inducing  
 CC principle (TIP) gene from Ti plasmid pTi15955. The sequence of the  
 CC T-DNA of the octopine-type Ti plasmid has fourteen open reading  
 CC frames bounded by eukaryotic promoters, ribosome binding sites, and  
 CC polyadenylation sites. With the vectors, expression of structural  
 CC foreign genes in plant cells is promoted. The gene esp. encodes an  
 CC insecticidal toxin identical to or derived from the crystal protein  
 CC of *Bacillus thuringiensis*.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 24596 BP; 6534 A; 5510 C; 5793 G; 6759 T; 0 other:  
 Query Match 98.0%; Score 1765.2; DB 6; Length 24596;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1792; Conservative 0; Mismatches 8; Indels 2; Gaps 2;  
 QY 1 ANGTCAGCTTCACTCTCTCTTATTAACGATGCGATCATCTCCCAACCAAAATGCTGAT 60  
 DB 5809 ATGTCAAGCTTCACTCTCTCTTATTAACGATGCGATCATCTCCCAACCAAAATGCTGAT 5868  
 QY 61 CTGACAAATGATGAAGCGGATGAATTGACCGGAGGTTTCGATGCTTTCTTAGAA 120  
 DB 5869 CTGACAAATGATGAAGCGGATGAATTGACCGGAGGTTTCGATGCTTTCTTAGAA 5928  
 QY 121 CGAAGAGCTTCTAGGGGAAAGAGATTACTCAAAGCTCCACGAGTGCAGCGCTGGGTTA 180  
 DB 5929 CGAAGAGCTTCTAGGGGAAAGAGATTACTCAAATCTCAACGAGTGCAGCGCTGGGTTA 5988  
 QY 181 GCTTGCAAAAGGCTGCGCATGCTGCTTCCCGGATTTCTCAGCTGCTGTAAGTAAGCA 240  
 DB 5989 GCTTGCAAAAGGCTGCGCATGCTGCTTCCCGGATTTCTCAGCTGCTGTAAGTAAGCA 6048  
 QY 241 GTTCTCTCGCTTATATATATTAATGCAAAATTTCTGGGCGGATCTTAATGAAA 300

```

Db 6049 GTTCTCTCCGCTTATATCTATATTGGCAAGAAATTCGCGGCGGATTAATTGATGAA 6108
Qy 301 CCTTGGGCGGGGCAACAGTAGTGTCTGTTGCCATGCACTTGACACCATTTGATG 360
Db 6109 CTTTGGGCGGGGCAACAGTAGTGTCTGTTGCCATGCACTTGACACCATTTGATG 6168
Qy 361 GATTTCTCCGAAACACAACATAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 420
Db 6169 GATTTCTCCGAAACACAACATAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 6228
Qy 421 CCGATTGATCTTATGATCACTTTCGTGGCCATTTCATCTTAAGACTGCGGCTTTCGAAC 480
Db 6229 CCGATTGATCTTATGATCACTTTCGTGGCCATTTCATCTTAAGACTGCGGCTTTCGAAC 6288
Qy 481 CTGCAATGCGGCTGTAACGGAATGGACAGATGAATAGCGTTACCGGGTTTACATTAAC 540
Db 6289 CTGCAATGCGGCTGTAACGGAATGGACAGATGAATAGCGTTACCGGGTTTACATTAAC 6348
Qy 541 CTGGAAGGGGCGGTGCACTTTCGATGATGATGATGATGATGATGATGATGATGATG 600
Db 6349 CTGGAAGGGGCGGTGCACTTTCGATGATGATGATGATGATGATGATGATGATGATG 6408
Qy 601 TCGGCAAGTTCTTTTCCAAACATGACTGCTCTAGCACTGCAAGCCGTTTTTGACCA 660
Db 6409 TCGGCAAGTTCTTTTCCAAACATGACTGCTCTAGCACTGCAAGCCGTTTTTGACCA 6468
Qy 661 TGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 6469 TGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6528
Qy 721 GTCAATTGCGCTGCGCACTTTCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 6529 GTCAATTGCGCTGCGCACTTTCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6588
Qy 781 GACGATGTTACAAATATATGAAAGCAAGTATGATGATGATGATGATGATGATGATG 840
Db 6589 GACGATGTTACAAATATATGAAAGCAAGTATGATGATGATGATGATGATGATGATG 6648
Qy 841 TTGAGGAGCGCTCTAGTGTCTGCGGCAATGGGGCGGATTCCTCCGCTGCA 900
Db 6649 TTGAGGAGCGCTCTAGTGTCTGCGGCAATGGGGCGGATTCCTCCGCTGCA 6708
Qy 901 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGCTTTCGATGAGCGGTTCCCAAT 960
Db 6709 TTCTGCTGTTTTTCTTCTCTGAGCGTTACGCGCTGCTTTCGATGAGCGGTTCCCAAT 6768
Qy 961 CCGGCAAGTGCACACTTACTTGCTTACCAAGCGCTTCCAAATCATGTGAAAAGCCGG 1020
Db 6769 CCGGCAAGTGCACACTTACTTGCTTACCAAGCGCTTCCAAATCATGTGAAAAGCCGG 6828
Qy 1021 CAGCTGCACCGAAGCTTTCATCGGTTTACCAAGGTTGGGCGTCTTGAAAGGAC 1080
Db 6829 CAGCTGCACCGAAGCTTTCATCGGTTTACCAAGGTTGGGCGTCTTGAAAGGAC 6888
Qy 1081 GGTTCATGAGGAGATATTTGTTGCTTCGCTGCTGCTTACTTACTCAGGCTTTGAAA 1140
Db 6889 GGTTCATGAGGAGATATTTGTTGCTTCGCTGCTGCTTACTTACTCAGGCTTTGAAA 6948
Qy 1141 TCAGACACATTAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACGTTTGGGAGG 1200
Db 6949 TCAGACACATTAGTGGGCTCATGACTCTGCGCAATTTGGCTGAACGTTTGGGAGG 7008
Qy 1201 GAGTCCCTCTCTCAGGAGTAGAGAGATCTTCTGGGCAACATCTCCTGGGCT-AA 1259
Db 7009 GAGTCCCTCTCTCAGGAGTAGAGAGATCTTCTGGGCAACATCTCCTGGGCTGAAA 7068
Qy 1260 ACATGAGTTTCTCTCATGATGGGACCTATTCAAGCTTAATGGGAATAGATCTGCGGG 1319
Db 7069 ACATGAGTTTCTCTCATGATGGGACCTATTCAAGCTTAATGGGAATAGATCTGCGG- 7127
Qy 1320 GTTTGCTCAGTTTAAAAAGCGGTTTATGAGATCTTCGCTTGTGTCATCAACGGATA 1379
Db 7128 GTTTGCTCAGTTTAAAAAGCGGTTTATGAGATCTTCGCTTGTGTCATCAACGGATA 7187

```

```

Qy 1380 TGAAGAAATCAGCGGATGTGCTCGAAGAAATCTCAGAACTTCCAGTCCGATC 1439
Db 7188 TGAAGAAATCAGCGGATGTGCTCGAAGAAATCTCAGAACTTCCAGTCCGATC 7247
Qy 1440 TGAAGTGTAAAGGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGCGAT 1499
Db 7248 TGAAGTGTAAAGGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGCGAT 7307
Qy 1500 TCAGAGGAAAAAGCAAAATTAAGATTAAGGCTTAAGACGGGATATCGAATTATGA 1559
Db 7308 TCAGAGGAAAAAGCAAAATTAAGATTAAGGCTTAAGACGGGATATCGAATTATGA 7367
Qy 1560 TAAAGTGTGTGTCATCTGCACTCGCAATATCCAACTCAGGATTTGCTGATGCGA 1619
Db 7368 TAAAGTGTGTGTCATCTGCACTCGCAATATCCAACTCAGGATTTGCTGATGCGA 7427
Qy 1620 TACCAATATTTTTCAGGACACAGTGAACCAAGCGTTGATTAACGCCATATGACAGATC 1679
Db 7428 TACCAATATTTTTCAGGACACAGTGAACCAAGCGTTGATTAACGCCATATGACAGATC 7487
Qy 1680 GTCAAAACCTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1739
Db 7488 GTCAAAACCTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7547
Qy 1740 TGTCTCATGAGACGGGATCGCAAAACAGTGTATTCCTGACTATGATGCGAGATCC 1799
Db 7548 TGTCTCATGAGACGGGATCGCAAAACAGTGTATTCCTGACTATGATGCGAGATCC 7607
Qy 1800 GA 1801
Db 7608 GA 7609

```

```

RESULT 5
ID AAN50226 standard; DNA; 24593 BP.
XX AAN50226;
AC AAN50226;
XX 25-MAR-2003 (updated)
DT 16-OCT-1991 (first entry)
XX
DE Sequence of opine synthase gene.
XX
KW Plant vector; Ti plasmid; T-DNA; ss.
XX
OS Agrobacterium tumefaciens Ti plasmid.
XX
PN EP140556-A.
XX
PD 08-MAY-1985.
XX
PF 12-SEP-1984; 84EP-0306233.
XX
PR 14-SEP-1983; 83US-0532280.
XX
PA (AGRK ) AGRIGENETICS RES ASSOC LTD.
PA (LUBR ) LUBRIZOL GENETICS INC.
XX
PI Dahl GA, Sutton DW, Barker RF;
XX
DR WPI; 1985-112088/19.
XX
PT Plasmid contg. opine synthase gene for selection - and foreign
PT DNA, useful as vector for transforming plant cells
XX
PS Disclosure; Page 212-217; 69pp; Japanese.
XX
CC Plant cells (and protoplasts) and plasmids contg. the DNA fragment
CC which includes an opine synthase gene plus a gene for antibiotic
CC resistance are claimed. These plasmids provide max. efficiency for
CC transfer of foreign genes and can be amplified in the plant genome.

```

CC They do not contain genes specifying tumour formation and will not  
CC spread antibiotic resistance throughout the plant population.  
CC (Updated on 25-MAR-2003 to correct PF field.)  
CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 24593 BP; 6554 A; 5493 C; 5787 G; 6759 T; 0 other;

Query Match 97.8%; Score 1762; DB 6; Length 24593;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1790; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

```
QY 1 ATGTACCTTACCTCTCTTATACAGTGCATCTCCCAACCAAAATGTGGAT 60
DB 5809 ATGTACCTTACCTCTCTTATACAGTGCATCTCCCAACCAAAATGTGGAT 5868
QY 61 CTGACATATGTCATTAAGGCGGATGAAATGGAACCGGCGGTTCCATGCTTTGAA 120
DB 5869 CTGACATATGTCATTAAGGCGGATGAAATGGAACCGGCGGTTCCATGCTTTGAA 5928
QY 121 CGAAGAGCTTCTAGGGGAGAGGATTAACAAGCTCAACGAGTGCAGCGGTGA 180
DB 5929 CGAAGAGCTTCTAGGGGAGAGGATTAACAAGCTCAACGAGTGCAGCGGTGA 5988
QY 181 GCTTGCAGGAGGCTGCGGATGCTGCTTCCGAGATCTCAGCTGTGGAAGTACGA 240
DB 5989 GCTTGCAGGAGGCTGCGGATGCTGCTTCCGAGATCTCAGCTGTGGAAGTACGA 6048
QY 241 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTCGGGCGGATCTTGAATGAAA 300
DB 6049 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTCGGGCGGATCTTGAATGAAA 6108
QY 301 CCTTGGGCGGCGGCAACAGTAGTGTCTGTTGCCATCGACTTGAACAATTTGGATG 360
DB 6109 CCTTGGGCGGCGGCAACAGTAGTGTCTGTTGCCATCGACTTGAACAATTTGGATG 6168
QY 361 GATTTCTCCGAGACAACTAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 420
DB 6169 GATTTCTCCGAGACAACTAATCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 6228
QY 421 CCGATTGATCTTATGTCATTTGCGGCAATTTCAATCTTAAAGTCGCGGCTTGAAC 480
DB 6229 CCGATTGATCTTATGTCATTTGCGGCAATTTCAATCTTAAAGTCGCGGCTTGAAC 6288
QY 481 CTCGCAATGCGGCTGATACGAGATGCGAGATGAATGCGTTACCGGTTTACATACC 540
DB 6289 CTCGCAATGCGGCTGATACGAGATGCGAGATGAATGCGTTACCGGTTTACATACC 6348
QY 541 CTTGAAGGGGCGGTGCATTTGACATGTAAGTTATGTTGGAACCTGATGCTGAAGGT 600
DB 6349 CTTGAAGGGGCGGTGCATTTGACATGTAAGTTATGTTGGAACCTGATGCTGAAGGT 6408
QY 601 TCGGCAAGTTCTTTCCAACTATGACTGCTCTACGACTGACGACCGTTTTTGAACAA 660
DB 6409 TCGGCAAGTTCTTTCCAACTATGACTGCTCTACGACTGACGACCGTTTTTGAACAA 6468
QY 661 TGTTCGATATGTCAGAGATGCGGCTTCTTCCGAGAGTTCCTTAAGCGGAAGTGGCG 720
DB 6469 TGTTCGATATGTCAGAGATGCGGCTTCTTCCGAGAGTTCCTTAAGCGGAAGTGGCG 6528
QY 721 GTTCATTGGCGCTGGCAATTTCCGAGCTGTTGTTGGAAGCAAGCTGCTTCAATGCT 780
DB 6529 GTTCATTGGCGCTGGCAATTTCCGAGCTGTTGTTGGAAGCAAGCTGCTTCAATGCT 6588
QY 781 GACGATGTTACAAATATATGAAAGCAAGTATGTTGTTGAGGCAAGCTTTGGTCAATGCT 840
DB 6589 GACGATGTTACAAATATATGAAAGCAAGTATGTTGTTGAGGCAAGCTTTGGTCAATGCT 6648
QY 841 TTCAGGAGGCTCTATGTCGTCGAGCGGAAATGGGGCGATATGCTTCTCTGCTGCA 900
DB 6649 TTCAGGAGGCTCTATGTCGTCGAGCGGAAATGGGGCGATATGCTTCTCTGCTGCA 6708
QY 901 TTCTGCTTGTTTTCTTCTCGAGCGTTACGCGCTTCTTCAATGAGGCGGTTCCCAAT 960
DB 901 TTCTGCTTGTTTTCTTCTCGAGCGTTACGCGCTTCTTCAATGAGGCGGTTCCCAAT 960
```

```
DB 6709 TTCTGCTTGTTTTCTTCTCGAGCGTTACGCGCTTCTTCAATGAGGCGGTTCCCAAT 6768
QY 961 CCGGACAGTGCACCTTATGTCCTACCAAGGCTCCAAATGTTGGAAGCCGG 1020
DB 6769 CCGGACAGTGCACCTTATGTCCTACCAAGGCTCCAAATGTTGGAAGCCGG 6828
QY 1021 CAGCTCCACCGAGCTGTTCCATCGGTTTCAACGTTGGCGCTGTTCTTGAAGAC 1080
DB 6829 CAGCTCCACCGAGCTGTTCCATCGGTTTCAACGTTGGCGCTGTTCTTGAAGAC 6888
QY 1081 GGTTCATGAGCGAGATATGTTGCTTCCGCTGCTGCTATTAATCAGGCTTGAAA 1140
DB 6889 GGTTCATGAGCGAGATATGTTGCTTCCGCTGCTGCTATTAATCAGGCTTGAAA 6948
QY 1141 TCAGACATATTAAGTGGGCTCATGCTCTTGGCAAAATTTGGCTGAACCGTTTGGGAGG 1200
DB 6949 TCAGACATATTAAGTGGGCTCATGCTCTTGGCAAAATTTGGCTGAACCGTTTGGGAGG 7008
QY 1201 GAGTCTTCTCTCAGGAGTAGAGAGATCTTCTGGGACACATCTCTGGTGGT-AA 1259
DB 7009 GAGTCTTCTCTCAGGAGTAGAGAGATCTTCTGGGACACATCTCTGGTGGTAA 7068
QY 1260 ACATGAGATTTTCTCATGATTTGGACCTATTCAAGCTAATGGGAATAGATCTGGCGG 1319
DB 7069 ACATGAGATTTTCTCATGATTTGGACCTATTCAAGCTAATGGGAATAGATCTGGC-GG 7127
QY 1320 GTTTGTCAGTTTGTGAAAGGGGTTATTTGAGATCTCTCCGCTTGTGATCAAGGAT 1379
DB 7128 GTTTGTCAGTTTGTGAAAGGGGTTATTTGAGATCTCTCCGCTTGTGATCAAGGAT 7187
QY 1380 TGAAGAAATTCAGCGGATGTCCTTGAAGAAATCTCAGAACTTCCAGTCCGATCGCATC 1439
DB 7188 TGAAGAAATTCAGCGGATGTCCTTGAAGAAATCTCAGAACTTCCAGTCCGATCGCATC 7247
QY 1440 TGAAGTGTAAAGTGTGTCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAGGCGCAT 1499
DB 7248 TGAAGTGTAAAGTGTGTCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAGGCGCAT 7307
QY 1500 TCAGAGGAAAGACAAAATTAAGTAAAGGCTTAAAGCGGGATATCTGAATCTTATGA 1559
DB 7308 TCAGAGGAAAGACAAAATTAAGTAAAGGCTTAAAGCGGGATATCTGAATCTTATGA 7367
QY 1560 TAAAGTGTGTGATCAGCTGCGCAATTCACAACTCAGGCAATGCTGCAATGCGA 1619
DB 7368 TAAAGTGTGTGATCAGCTGCGCAATTCACAACTCAGGCAATGCTGCAATGCGA 7427
QY 1620 TACCAATATTTTTCAGGACACAGTGAACCAAGCGTTGATTAACGACATATGACAGATC 1679
DB 7428 TACCAATATTTTTCAGGACACAGTGAACCAAGCGTTGATTAACGACATATGACAGATC 7487
QY 1680 GTCAAAATCTTCTGATGATGTAACGAAATTTCTGTTTGAACATATCTCCGCTTGG 1739
DB 7488 GTCAAAATCTTCTGATGATGTAACGAAATTTCTGTTTGAACATATCTCCGCTTGG 7547
QY 1740 TGTTCCTCATGAGCGGATGCAAAAGCAAGTATGCTGAGCTATGATCCAGAGATCC 1799
DB 7548 TGTTCCTCATGAGCGGATGCAAAAGCAAGTATGCTGAGCTATGATCCAGAGATCC 7607
QY 1800 GA 1801
DB 7608 GA 7609
```

RESULT 6  
AAV69756  
ID AAV69756 standard; DNA; 2211 BP.  
XX  
XX AAV69756;  
XX  
XX 01-MAR-1999 (first entry)  
XX  
XX Tryptophan oxygenase iaam gene.  
DE

KM Tryptophan oxygenase; *iaam* gene; auxin; indoleacetic acid;  
KM transgenic plant; seedless fruit; tomato; watermelon; cucumber; da.  
XX  
OS Agrobacterium tumefaciens.  
XX  
FH Key Location/Qualifiers  
FT CDS 4..2205  
FT /\*tag= a  
XX  
XX MO9849888-A1.  
XX  
XX 12-NOV-1998.  
XX  
PD 06-MAY-1998; 98MO-US09013.  
XX  
PF 06-MAY-1997; 97US-0045725.  
XX  
PR 06-MAY-1997; 97US-0045725.  
XX  
PA (UNIV ) UNIV KANSAS STATE RES FOUND.  
XX  
PI L4 Y;  
XX  
XX WPI; 1999-034673/03.  
XX  
XX P-PSDB; AAW81574.  
XX  
XX A new construct to express phytohormones in developing fruit -  
PT useful for, e.g. producing substantially seedless fruit from  
PT transgenic plants  
XX  
XX Example 1; Page 21-26; 49pp; English.  
XX  
XX This is the nucleotide sequence of the Agrobacterium tumefaciens  
XX *iaam* gene that codes for tryptophan oxygenase (see AAW81574). This  
XX enzyme catalyses the conversion of tryptophan to indoleacetic acid  
XX in the biosynthetic pathway of the auxin indoleacetic acid. A  
XX claimed DNA construct comprises either an isopentenyl transferase  
XX (see AAW81575) or a tryptophan oxygenase encoding sequence, operably  
XX linked to an ovary or developing fruit-specific plant-expressible  
XX promoter (see AAW69755 and AAW69759). The construct is used to stably  
XX integrate enzymes involved in cytokinin or auxin biosynthesis into  
XX the plant genome to achieve a transgenic plant (preferably tomato,  
XX cucumber or watermelon) producing seedless fruit in the absence of  
XX pollination.  
SQ Sequence 2211 BP; 528 A; 504 C; 595 G; 584 T; 0 other;  
Query Match 94.2%; Score 1697.4; DB 20; Length 2211;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1726; Conservative 0; Mismatches 11; Indels 2; Gaps 2;  
QY 64 ACAATGTCGATTAAGGCGATGTAATGACGCGAGGGTTCCGATGCTTCTTAGAACA 123  
Db 1 ACAATGTCGATTAAGGCGATGTAATGACGCGAGGGTTCCGATGCTTCTTAGAACA 60  
QY 124 GAAGCTTCTAGGGAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGGTTAGCT 183  
Db 61 GAAGCTTCTAGGGAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGGTTAGCT 120  
QY 184 TGCAGAAAGGCTGGCCGATGCTGCTTCCCGAGATCTCAGCTGTGAAAGTACGATT 243  
Db 121 TGCAGAAAGGCTGGCCGATGCTGCTTCCCGAGATCTCAGCTGTGAAAGTACGATT 180  
QY 244 CTCCTCCCTATATCTATATTTGGCAAGAAATTCGGGGCGGATTAATTGAATGAAACCT 303  
Db 181 CTCCTCCCTATATCTATATTTGGCAAGAAATTCGGGGCGGATTAATTGAATGAAACCT 240  
QY 304 TGGGCGGCGGCAACAGTGAAGTCTCTGTCGCAATTCGACATTTGGATGATGAT 363  
Db 241 TGGGCGGCGGCAACAGTGAAGTCTCTGTCGCAATTCGACATTTGGATGATGAT 300  
QY 364 TTCTCCGAGACCAACTAATTCAGAGCCCTGTTTGTCTGAGCGGTAAAGATGACCG 423  
Db 301 TTCTCCGAGACCAACTAATTCAGAGCCCTGTTTGTCTGAGCGGTAAAGATGACCG 360

QY 424 ATTGATCTTAGCATTTCCGCGCATTTTCATCTTAAGACTGCCGCTTTCGAACCTG 483  
Db 361 ATTGATCTTAGCATTTCCGCGCATTTTCATCTTAAGACTGCCGCTTTCGAACCTG 420  
QY 484 CCAATGCCCTGTAGCAGAAATGACAGATGAATGCGTTTACCGGGTTTACCAATACCCTT 543  
Db 421 CCAATGCCCTGTAGCAGAAATGACAGATGAATGCGTTTACCGGGTTTACCAATACCCTT 480  
QY 544 GAAAGGCGCGTGCATTTGACATGATGATGCTTATGTCGAAACCTGATGTCGAAGGGTTG 603  
Db 481 GAAAGGCGCGTGCATTTGACATGATGATGCTTATGTCGAAACCTGATGTCGAAGGGTTG 540  
QY 604 GCAGTTTCTTTCCAAACAATGCACTTCTTACGATGCAAGCCGTTTTTTCACATGT 663  
Db 541 GCAGTTTCTTTCCAAACAATGCACTTCTTACGATGCAAGCCGTTTTTTCACATGT 600  
QY 664 TCCGATAGTGAACGATGCGCTTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTC 723  
Db 601 TCCGATAGTGAACGATGCGCTTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTC 660  
QY 724 ATTGGGCGTGGCATTTCCGGAATCGTGTGGCAAGAACTGCTTATGCTGGGGTAGAC 783  
Db 661 ATTGGGCGTGGCATTTCCGGAATCGTGTGGCAAGAACTGCTTATGCTGGGGTAGAC 720  
QY 784 GATGTTACAATATATAGAGCAAGTGAATGATGCTGTGAGGCAAGCTTGTGCATGCTTTC 843  
Db 721 GATGTTACAATATATAGAGCAAGTGAATGATGCTGTGAGGCAAGCTTGTGCATGCTTTC 780  
QY 844 AGGAGCGCTCTTAATGCTGCTGCGCAATGGGGCGATGCGATTTCTCTGCTGCATTC 903  
Db 781 AGGAGCGCTCTTAATGCTGCTGCGCAATGGGGCGATGCGATTTCTCTGCTGCATTC 840  
QY 904 TGCTTTTCTTCTCTCGAGGTTACGGGCTTCTTTCGATGAGGCGGTTCCCAATATCC 963  
Db 841 TGCTTTTCTTCTCTCGAGGTTACGGGCTTCTTTCGATGAGGCGGTTCCCAATATCC 900  
QY 964 GGCACAGTGCACACTTACTTGTGTCAACAAGGCTCAATACATGTGAAGCCGGGAC 1023  
Db 901 GGCACAGTGCACACTTACTTGTGTCAACAAGGCTCAATACATGTGAAGCCGGGAC 960  
QY 1024 CTGCCACGAAAGTGTTCATCGCGTTTACACAGGTTGGGCTGCTTCTTGAAGACGCT 1083  
Db 961 CTGCCACGAAAGTGTTCATCGCGTTTACACAGGTTGGGCTGCTTCTTGAAGACGCT 1020  
QY 1084 TTTCATGAGGAGATATGTTGTTGGCTGCGCTGCTGCTATTAATCAAGGCTTGAATCA 1143  
Db 1021 TTTCATGAGGAGATATGTTGTTGGCTGCGCTGCTGCTATTAATCAAGGCTTGAATCA 1080  
QY 1144 GGAACATTAAGTGGGCTATGACTCTCTGCAAAATTTGGCTGAACCGTTTCGGAGGGAG 1203  
Db 1081 GGAACATTAAGTGGGCTATGACTCTCTGCAAAATTTGGCTGAACCGTTTCGGAGGGAG 1140  
QY 1204 TCTTCTCTTACAGGATTAAGAGGATCTTCTGGGACACATCTCTCTGTGTGT -AACA 1262  
Db 1141 TCTTCTCTTACAGGATTAAGAGGATCTTCTGGGACACATCTCTCTGTGTGTGAACA 1200  
QY 1263 TGGAGTTTCTCATATATGAGGACTTATTCAGACTAATGGAATTAAGATTTGGC -GGGTT 1322  
Db 1201 TGGAGTTTCTCATATATGAGGACTTATTCAGACTAATGGAATTAAGATTTGGC -GGGTT 1259  
QY 1323 TGGTCCAGTTTGAAGAGGGGTTTATTAAGATCTCTCGCTTGGTCAACAACGATATGA 1382  
Db 1260 TGGTCCAGTTTGAAGAGGGGTTTATTAAGATCTCTCGCTTGGTCAACAACGATATGA 1319  
QY 1383 AGAAATACGCGGATGTGCTTGAAGGAATCTAGAACTTCCAGCTGCGATGCGATCTGA 1442  
Db 1320 AGAAATACGCGGATGTGCTTGAAGGAATCTAGAACTTCCAGCTGCGATGCGATCTGA 1379  
QY 1443 AGTGTTAACGCGTGTCTGTGAGCCAGGCAATATGCAATGTTCAAGTCAAGGCGATTCA 1502  
Db 1380 AGTGTTAACGCGTGTCTGTGAGCCAGGCAATATGCAATGTTCAAGTCAAGGCGATTCA 1439  
QY 1503 GAAGGAAAGACAAAATAAAGATTAAGGCTTAAGAGCGGGATATCTGAACCTTATGATTA 1562

DB 1440 GAAGGAAAAAACAATAAAGATTAAGCGGATATCTGAATTATGATTA 1499  
QY 1563 GGTGGGTGTCACATCTGGACTCCGAATATCCAACTAGCATTTGCTGACATGCCATAC 1622  
DB 1500 GGTGGGTGTCACATCTGGACTCCGAATATCCAACTAGCATTTGCTGACATGCCATAC 1559  
QY 1623 CAATATTTTTCAGGACCACTGAAACCAAGCGGTGATTAACGCAATATGACAGATGCTC 1682  
DB 1560 CAATATTTTTCAGGACCACTGAAACCAAGCGGTGATTAACGCAATATGACAGATGCTC 1619  
QY 1683 AAAACTCTTCTGATGATCTGACGAAATTTCTGGTTAGACATATCTCCGCTTGTGT 1742  
DB 1620 AAAACTCTTCTGATGATCTGACGAAATTTCTGGTTAGACATATCTCCGCTTGTGT 1679  
QY 1743 CCTCATGAGAGGATCGCAAAAGCAGTGTATTCCTGAGACTATGATGCGAGATCCGA 1801  
DB 1680 CCTCATGAGAGGATCGCAAAAGCAGTGTATTCCTGAGACTATGAGCCGAGATCCGA 1738

## RESULT 7

AAT48868/C  
ID AAT48868 standard; DNA; 6112 BP.

AC AAT48868;  
XX  
DT 07-APR-1997 (first entry)

DE Iams expression vector pPH5838.

XX Male sterile plant; hybrid seed; pollen; microspore; oilseed;

KW canola; rape; antisense RNA; stress resistance;

KW herbicide resistance; transgenic plant; crop protection;

KW Iams; indole acetamide synthase; Bp10; promoter; Brassica;

KW pPH5838; ds; cyclic.

XX

OS Chimeric Brassica sp.;

OS Chimeric Agrobacterium tumefaciens.

XX

PH Key Location/Qualifiers

FT misc\_difference 3794 /tag= a

FT /note= "Base 3794 is given as v in the specification"

XX

PN WO9640950-A1.

XX 19-DEC-1996.

PD 07-JUN-1996; 96WO-US08692.

PF 07-JUN-1995; 95US-0476864.

PR (PION-) PIONEER HI-BRED INT INC.

XX

PA Albani DJ, Arnison PG, Fabijanski SF, Laurian RS;

PI WPI; 1997-052340/05.

XX

DR Recombinant molecule for preparation of male sterile plants - used

XX to produce stress resistant plants and for increase in seed yield

PT Example 39a; Fig 27a-e; 224pp; English.

XX

PS

XX

CC plasmid pPH5838 (AAT48868) contains the Brassica microspore-specific

CC Bp10 promoter linked to the indole acetamide synthase (Iams) gene

CC from the Agrobacterium tumefaciens Ti plasmid derivative pPCV311.

CC and to the nos 3' non-translated region. The Iams gene is involved

CC in tumour formation in Agrobacterium-infected plant. Plasmids

CC pPH5838-41 (see also AAT48869-71), contg. Iams or Iams genes under

CC control of another-specific Bp10 or TA39 promoters, were introduced

CC into canola lines by transformation. Male sterile plants were

CC obtd. These can be used in the prodn. of hybrid seed using pollen

CC from a male fertile line.

XX Sequence 6112 BP; 1515 A; 1466 C; 1451 G; 1679 T; 1 other;

SO Query Match 88.9%; Score 1600.4; DB 18; Length 6112;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1689; Conservative 0; Mismatches 111; Indels 2; Gaps 2;

QY 1 ATGTGACCTTCACTCTCTTGTATTAACAGTGGATCATCTCCCAACCAAAATGATGAT 60  
DB ATGTGACCTTCACTCTCTTGTATTAACAGTGGATCATCTCTCCCAAAATGATGAT 2811  
QY 61 CTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB CTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2751  
QY 121 CGAAGACTTCTAGGGAAGAGAGATTAACAAGCTCCACCGAGTGCAGCGTGGTTA 180  
DB CGTGAAGCTTCTAGGGAAGAGAGATTAACAAGCTCCACCGAGTGCAGCGTGGTTA 2691  
QY 181 GCTTGCAGAAAGCTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
DB GCTTGCAGAAAGCTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2631  
QY 241 GTTCTCTCGCTTATATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
DB GCTTCTCTCGCTTATATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2571  
QY 301 CCTTGGCGCGGCGCAAGTGAAGTGTCTGTTGCCATGCACTTGGACATTTTGCATG 360  
DB CCTTGGCGCGGCGCAAGTGAAGTGTCTGTTGCCATGCACTTGGACATTTTGCATG 2511  
QY 2570 CCTTGGCGCGGCGCAAGTGAAGTGTCTGTTGCCATGCACTTGGACATTTTGCATG 2511  
QY 361 GATTTTCTCGAAGACACATTAATCCAGCCGCTTTTGTGAGCGGTAAGATGCA 420  
DB GATTTTCTCGAAGACACATTAATCCAGCCGCTTTTGTGAGCGGTAAGATGCA 2451  
QY 421 CGATGATCTTATGATTTTGTGAGCGGCTTGTGAGCGGCTTGTGAGCGGCTTGTGAGCG 480  
DB TCGAGGATCTTATGATTTTGTGAGCGGCTTGTGAGCGGCTTGTGAGCGGCTTGTGAGCG 2291  
QY 481 CTGCAATGCGCGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 540  
DB CTGCAATGCGCGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2231  
QY 2390 CTGCAATGCGCGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2231  
DB CTGCAATGCGCGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2231  
QY 541 CTGCAATGCGCGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 600  
DB CTGCAATGCGCGCTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2271  
QY 601 TCGGCAAGTTCCTTCCCAATGCACTGCTCTAGATGCAAGCGTTTGTGACCA 660  
DB TCGGCAAGTTCCTTCCCAATGCACTGCTCTAGATGCAAGCGTTTGTGACCA 2211  
QY 661 TGTTCGATAGTGAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 720  
DB TGTTCGATAGTGAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2211  
QY 2210 TGTTCGATAGTGAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2151  
DB TGTTCGATAGTGAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2151  
QY 721 GTTATTTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
DB GTTATTTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2150  
QY 781 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
DB GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2091  
QY 2090 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2031  
DB GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2031  
QY 841 TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1971  
QY 2030 TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1971  
DB TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1971  
QY 901 TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1911  
QY 1970 TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1911  
DB TTCAGGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1911  
QY 961 CCCGCAAGTGCACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
DB CCCGCAAGTGCACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020

Db 1910 CCGGGACAGTCCGACACTAATTGCTTACCAAGGCTCCGATACATGTGGAAAGCCGGG 1851  
 Qy 1021 CAGCTGCCACCGAAAGCTGCTTCATCCGGTTTACAAGGTTGGCGGCTTGAAGAC 1080  
 Db 1850 CACGAGCCACCGAAGCTGTTCCATCCGGTTTACAGCCGTTGGCGGCTTGAAGAC 1791  
 Qy 1081 GATTTTCATAGGAGATATTGTTGGCTCCGCTGCTATTAATCTAGAGCCCTTAAA 1140  
 Db 1790 GATTTTCATAGGAGATATTGTTGGCTCCGCTGCTATTAATCTAGAGCCCTTAAA 1731  
 Qy 1141 TCAGACACATTAAGTGGGCTCATGACTCTGCGAAATTTGGCTGAAACGTTTGGAGG 1200  
 Db 1730 TCAGGAGACATTAAGCGGGCTCATGACTCTGCGAAATTTGGCTGAAACGTTTGGAGG 1671  
 Qy 1201 GAGTCTCTTCTTCAAGGATAGAGAGATCTTTTGGGACACATCTCTCGTGGT-AA 1259  
 Db 1670 GAGTCTCTTCTTCAAGGATAGAGAGATCTTTTGGGACACATCTCTCGTGGT-AA 1611  
 Qy 1260 ACATGAGTTTCTCTCATGATTGGGACCTATTCAGATATGGGATAGGATCTGGGGG 1319  
 Db 1610 ACATGAGTTTCTCTCATGATTGGGACCTATTCAGATATGGGATAGGATCTGGGGG 1552  
 Qy 1320 GTTGTGTCAGTTTGAAGCGGGTTTATTAAGATCTCCGCTGCTCATCAAGGATA 1379  
 Db 1551 GTTGTGTCAGTTTGAAGCGGGTTTATTAAGATCTCCGCTGCTCATCAAGGATA 1492  
 Qy 1380 TGAAGAAATCAGCGGATGTCCTTGAAGGATCTCAGAACTTCCAGTGGATAGCCAC 1439  
 Db 1491 TGAAGAAATCAGCGGATGTCCTTGAAGGATCTCAGAACTTCCAGTGGATAGCCAC 1432  
 Qy 1440 TGAAGTGTAAAGGCTGCTGTCGAGCCAGCGCATATGCGATGTCAGATCAGGGGAT 1499  
 Db 1431 TCAAGTGTAAAGGCTGCTGTCGAGCCAGCGCATATGCGATGTCAGATCAGGGGAT 1372  
 Qy 1500 TCAGAGAGAAAGACAAATAAAGATTAAGCTTAAGAGCGGATATCTGAATTTATGA 1559  
 Db 1371 TGAGAAAGAAAGACAAATAAAGATTAAGCTTAAGAGCGGATATCTGAATTTATGA 1312  
 Qy 1560 TAAAGTGTGTGTCATCTGAGCTCGCAATATCAACTCAGGATGCTGACATCGGA 1619  
 Db 1311 TAAAGTGTGTGTCATCTGAGCTCGCAATATCAACTCAGGATGCTGACATCGGA 1252  
 Qy 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATPAAGCCATATGACAGATC 1679  
 Db 1251 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATPAAGCCATATGACAGATC 1192  
 Qy 1680 GTCAAAACTCTTCTGATGACTGAACGAAATTTGTTAGACCATATCTCCGCTTG 1739  
 Db 1191 GTCAAAACTCTTCTGATGACTGAACGAAATTTGTTAGACCATATCTCCGCTTG 1132  
 Qy 1740 TGTCTCATGAGCGGATTCGCAAAAGCAAGTATGCTGAGCTATAGTGTGCGAGATCC 1799  
 Db 1131 TGTCTCATGAGCGGATTCGCAAAAGCAAGTATGCTGAGCTATAGTGTGCGAGATCC 1072  
 Qy 1800 GA 1801  
 Db 1071 GA 1070

## RESULT 8

AAT48870/c  
 AAT48870 standard; DNA; 7208 BP.

AC AAT48870;  
 AT 07-APR-1997 (first entry)  
 DE 1amh expression vector pPH5840.  
 KM Male sterile plant; hybrid seed; pollen; microspore; oilseed;  
 KM canola; rape; antisense RNA; stress resistance;  
 KM herbicide resistance; transgenic plant; crop protection;

KM 1amh; indole acetamide hydrolase; TA39; promoter; tobacco;  
 KM pPH5840; ds; cyclic.  
 OS Chimeric Nicotiana tabacum;  
 OS Chimeric Agrobacterium tumefaciens.  
 FH Key Location/Qualifiers  
 FT misc\_difference 4890  
 FT /\*tag= a  
 FT /note= "base 4890 is given as v in the specification"

MO640950-A1.  
 19-DEC-1996.  
 07-JUN-1996; 96WO-US08692.  
 07-JUN-1995; 95US-0476864.  
 (PION-) PIONEER HI-BRED INT INC.  
 Albani DJ, Arnison PG, Fabijanski SF, Laurian RS;  
 WPI; 1997-052340/05.  
 Recombinant molecule for preparation of male sterile plants - used  
 PT to produce stress resistant plants and for increase in seed yield  
 PS Example 39a; Fig 29a-f; 224p; English.  
 CC Plasmid pPH5840 (AAT48870) contains the tobacco anther-specific  
 CC TA39(8B3) promoter and the 0' nontranslated leader operably linked  
 CC to the indole acetamide hydrolase (iamh) gene from the Agrobacterium  
 CC tumefaciens Ti plasmid derivative pCV311, and to the P1-II 3'  
 CC nontranslated region. The iamh gene is involved in tumour  
 CC formation in Agrobacterium-infected plants. Plasmids pPH5838-41  
 CC (see also AAT48868-71), contg. Iams or iamh genes under control of  
 CC anther-specific Bp10 or TA39 promoters, were introduced into canola  
 CC lines by transformation. Male sterile plants were obtd. These can  
 CC be used in the prodn. of hybrid seed using pollen from a male  
 CC fertile line.  
 CC XX

SQ Sequence 7208 BP; 1938 A; 1619 C; 1640 G; 2010 T; 1 other;  
 Query Match 88.9%; Score 1600.4; DB 18; Length 7208;  
 Best Local Similarity 93.7%; Pred. No. 0;  
 Matches 1689; Conservative 0; Mismatches 111; Indels 2; Gaps 2;

Qy 1 ATGTCAGCTTCACTCTCTTGAATTAACAGTGGATCATCTCCCAACCAATGATGAT 60  
 Db 2924 ATGTCAGCTTCACTCTCTTGAATTAACAGTGGATCATCTCTCCCAACCAATGATGAT 2865  
 Qy 61 CTGACATAGTGTGATTAAGCGGATGAATTGACCGCAGGGTTTCCAGTGCCTTTAGAA 120  
 Db 2864 CTGATATAGTGTGATTAAGCGGATGAATTGACCGCAGGGTTTCCAGTGCCTTTAGAA 2805  
 Qy 121 CGAAGAGCTTCTAGGGGAAGAGATTAATCAAAAGCTCCACCGAGTGCACGGCTGGTTA 180  
 Db 2804 CGTGAAGCTTCTAGGGGAAGAGATTAATCAAAAGCTCCCGGCGAGTGCACGGCTGGTTA 2745  
 Qy 181 GCTTGAAGAGGCTGGCGGATGATGCTGCTCCGAGATCTCAGTGTGAGAAAGTGA 240  
 Db 2744 GCTTGAAGAGGCTGGCGGATGATGCTGCTCCGAGATCTCAGTGTGAGAAAGTGA 2685  
 Qy 241 GTTCTCTCCGCTTATATCTATTTATTTGGCAAGAAATTTGGGGCGGATTAATCGAAA 300  
 Db 2684 GCCTCTCGGCTTATATCTATTTATTTGGCAAGAAATTTGGGGCGGATTAATCGAAA 2625  
 Qy 301 CTTTGGGCGGCGACAGTGAAGTGTCTGTTGCCATTCGACTTGGACACACATTTGCATG 360  
 Db 2624 CTTTGGGCGGCGACAGTGAAGTGTCTGTTGCCATTCGACTTGGACACACATTTGTATG 2565

QY 361 GATTTCGAGACACAACTAATCAAGCCCTGTTTGTGAGCGGTAAAGATGCA 420  
Db 2564 GATTTCGAGACACAACTTCTCAACCTGTTTGTGAGCGGTAAAGATGCA 2505  
QY 421 CCGATTGATTTAGTCACTTTGTGGCATTTCATCTGAAGCTGCGGCTTTGAAAC 480  
Db 2504 TCAGGAGATCTTGTGATTTGTGGCATTTCATCTGAAGCTGCGGCTTTGAAAC 2445  
QY 481 TCAGGAGATCTTGTGATTTGTGGCATTTCATCTGAAGCTGCGGCTTTGAAAC 540  
Db 2444 TCAGGAGATCTTGTGATTTGTGGCATTTCATCTGAAGCTGCGGCTTTGAAAC 2385  
QY 541 CTTGAAGGGGCGTGCATTTGACATGATGTTAGTGAAGCTGATGCTGAAGGT 600  
Db 2384 CTTGAAGGGGCGTGCATTTGACATGATGTTAGTGAAGCTGATGCTGAAGGT 2325  
QY 601 TCAGGAGATCTTGTGATTTGTGGCATTTCATCTGAAGCTGCGGCTTTGAAAC 660  
Db 2324 TCAGGAGATCTTGTGATTTGTGGCATTTCATCTGAAGCTGCGGCTTTGAAAC 2265  
QY 661 TGTTCGATGATGAGGAGATGCGGCTTTCCGAGAGATTTCTTAAGCCGAAAGTGGC 720  
Db 2264 TGTTCGATGATGAGGAGATGCGGCTTTCCGAGAGATTTCTTAAGCCGAAAGTGGC 2205  
QY 721 GTCAATTGCGCTGCACTTTCGAGACTGATGATGAGCAAGCTGCTTCACTGAGGATA 780  
Db 2204 ATCATTTGCGCTGCACTTTCGAGACTGATGATGAGCAAGCTGCTTCACTGAGGATA 2145  
QY 781 GACGATGTTAACAATATATGAAGCAAGTATGATGATGAGCAAGCTTGTGACATGCT 840  
Db 2144 GACGATGTTAACAATATATGAAGCAAGTATGATGATGAGCAAGCTTGTGACATGCT 2085  
QY 841 TTCAGGAGCGCTCTAGTGTGTCGTCGCGCAAAATGGGCGCATGCTGCTCTGCTCA 900  
Db 2084 TTCAGGAGCGCTCTAGTGTGTCGTCGCGCAAAATGGGCGCATGCTGCTCTGCTCA 2025  
QY 901 TTTCTGCTGTTTCTTCTCTGAGCGCTTACGCGCTGCTGATGAGCGCTTCCCAAT 960  
Db 2024 TTTCTGCTGTTTCTTCTCTGAGCGCTTACGCGCTGCTGATGAGCGCTTCCCAAT 1965  
QY 961 CCGGAGACAGTGCACCTTACTGATGATGATGAGCGCTTCCCAATGATGAGCGCGG 1020  
Db 1964 CCGGAGACAGTGCACCTTACTGATGATGATGAGCGCTTCCCAATGATGAGCGCGG 1905  
QY 1021 CAGCTCCACCGAAGCTTCCATCGGCTTACAGCGTTGGCGTCTTGAAGAC 1080  
Db 1904 CAGCTCCACCGAAGCTTCCATCGGCTTACAGCGTTGGCGTCTTGAAGAC 1845  
QY 1081 GGTTCATGAGGAGATATGTTGGTGGCTGCGCTGCTGATGATGAGCGCTTGAAG 1140  
Db 1844 GGTTCATGAGGAGATATGTTGGTGGCTGCGCTGCTGATGATGAGCGCTTGAAG 1785  
QY 1141 TCAGGACATTAAGTGGCTCATGATCTCGGCAATTTGCTGAACGCTTGGGAGG 1200  
Db 1784 TCAGGACATTAAGTGGCTCATGATCTCGGCAATTTGCTGAACGCTTGGGAGG 1725  
QY 1201 GAGTCTTCTTCTGAGGATGAGAGATCTTCTGGGCAACATCTCTCTGTGT- AA 1259  
Db 1724 GAGTCTTCTTCTGAGGATGAGAGATCTTCTGGGCAACATCTCTCTGTGTGTAA 1665  
QY 1260 ACATGAGTTTCTCATGATTTGGGACCTATTGAAGCTTAAGGATTTGGCGGG 1319  
Db 1664 ACATGAGTTTCTCATGATTTGGGACCTATTGAAGCTTAAGGATTTGGCGGG 1606  
QY 1320 GTTTCGATGTTTGAAGCGGTTTATGAGATCTCGGCTGATGATCAACGGATA 1379  
Db 1605 GTTTCGATGTTTGAAGCGGTTTATGAGATCTCGGCTGATGATCAACGGATA 1546  
QY 1380 TGAAGAAATCAGCGATGCTCTGAAGAAATCTGAAGCTTCAAGCTGATGCAATC 1439  
Db 1545 TGAAGAAATCAGCGATGCTCTGAAGAAATCTGAAGCTTCAAGCTGATGCAATC 1486  
QY 1440 TGAAGTGTAAACGGTGTCTGTGAGCAAGCGCATATGCAATGTTCAAGTCAAGCGCAT 1499

Db 1485 TCAGTGTGTTAAGCGTGTGCTGTTAAGCCAGCTATACGCAATGTTCAAGTCAAGCGGAT 1426  
QY 1500 TCAGAAAGAAAGACAAAATTAAGATTAAGCTTAAGAGCGGATATCTGAATTTATGA 1559  
Db 1425 TCAGAAAGAAAGACAAAATTAAGATTAAGCTTAAGAGCGGATATCTGAATTTATGA 1366  
QY 1560 TAAGTGTGTGTCATCTGAGTCTGCAATATCCAACTCAGGCAATTCCTGATGCGGA 1619  
Db 1365 TAAGTGTGTGTCATCTGAGTCTGCAATATCCAACTCAGGCAATTCCTGATGCGGA 1306  
QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGTTGATTAACAGCATATGACAGATC 1679  
Db 1305 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGTTGATTAACAGCATATGACAGATC 1246  
QY 1680 GTCAAAATCTTCTCTGATGATGATGAGCAAGAAATTTGTTAGACATATCTCCGTTTG 1739  
Db 1245 GTCAAAATCTTCTCTGATGATGATGAGCAAGAAATTTGTTAGACATATCTCCGTTTG 1186  
QY 1740 TGTTCATGATGAGCGGATGCGAAAGCAAGTATGCTGAGTATGAGTGGAGATCC 1799  
Db 1185 TGTTCATGATGAGCGGATGCGAAAGCAAGTATGCTGAGTATGAGTGGAGATCC 1126  
QY 1800 GA 1801  
Db 1125 GA 1124

RESULT 9  
AAN60970  
ID AAN60970 standard; DNA; 21126 BP.  
XX  
AC AAN60970;  
XX  
DT 25-MAR-2003 (updated)  
DT 28-OCT-1991 (first entry)  
XX  
DE TL-DNA region of A. rhizogene agropine-type plasmid pRiR1.  
XX  
KW Ri plasmid; plant plasmid.  
XX  
OS Agrobacterium rhizogenes HRI and A4.  
XX  
PN BP204590-A.  
XX  
PD 10-DEC-1986.  
XX  
PF 21-APR-1986; 86EP-0400855.  
XX  
PR 22-APR-1985; 85US-0725368.  
XX  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX  
PI Slighton JL, Tepfer DA;  
XX  
DR WPI, 1986-327378/50.  
XX  
PT Plant cell genetic modification - using pRi T-DNA promoter to  
PT control expression of heterologous foreign structure genes, e.g.  
PT for phaseolin.  
XX  
PS Disclosure; Fig 2; 86EP; English.  
XX  
CC The TL-DNA sequence is that of an Agrobacterium Ri plasmid, it  
CC contains sixteen open reading frames, bounded by eukaryotic  
CC promoters, ribosome binding and polyadenylation site. The plasmid  
CC sequence may be used for genetic manipulation of plants, conveying  
CC phenotypes such as improved resistance to environmental stress, disease  
CC and parasitism.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 21126 BP; 5359 A; 4901 C; 5308 G; 5558 T; 0 other;

Query Match	6.8%	Score 121.6	DB 7	Length 21126
Best Local Similarity	45.7%	Pred. No. 5.7e-28		
Matches	540	Conservative 0	Mismatches 634	Indels 8
Gaps				3
Qy	623	TCGACTTGCTCTACGACTGACAGCCGTTTCTTGAACAATGTTCCGATAGTGGACGATCG	682	
Db	7273	TTGACCTCGCGGTACGACTATCAGGTTTTCTCCAGAGAGTGGATGCCCATGATGCTGAG	7332	
Qy	683	GCTTCTTTCCGGAGAGATGTTCTTAAGCCGAAAGTGGCGGTATTTGGCGTGGCATTTCCG	742	
Db	7333	GGCATTTATCCGGAAAGGCGCACCACTACTTATGTTGCCATGTGAGAGCGGGCTGCTG	7392	
Qy	743	GACTCGATGGTGGCAACGAACTGCTTCAATGCTGGGGTGAAGATGTTCAATATATATGAAG	802	
Db	7393	GCTTTGTTGCTCCACAGACTACTTTGGCGCTGGCGCTAGAGAAATCACTCTTTTCGATA	7452	
Qy	803	CAAGTGAATCGGT--TGAGGCAAGCTTTGGTCACATGCTTTGAGGAGCGCTCTAGTG	859	
Db	7453	CCGTTGATGAGATCCGATGTTTGGGGGATGCGCGATGCCAAAGGGAGGCTCACAGG	7512	
Qy	860	TCGTGGCCGAAATGGGGCGATGGCATTTCTCTGCTGCATTTCTGTTGTTTTCTTCC	919	
Db	7513	CCTTGAGCTCGTTGCGTGTCAATGCTTTCTCCGCAACCAATTTGCTGCTGATACATAC	7572	
Qy	920	TCGAGCGTTAGGCGCTGTCTCGATGAGCGCGTCCCAATCCGGGCACTGCACACTT	979	
Db	7573	TGGATTAAGTTTAAATATCCGTCAGCCTTCTTTCTTGCCGGACGACACACACAG	7632	
Qy	980	ACTTGCTACCAAGCGTCCCAATACATGATGAAAGCCGGGCGAGCTGCCACCGAAGCTGT	1039	
Db	7633	CACATATATTTCCGCAAGAAAGCTATCGATGGAAGCGGGGCAAGCTCCGCGGGGATAT	7692	
Qy	1040	TCCATTCGCGTTTACACGCGTTGGCGTGTCTTGAAGACGTTTTCATAGCGAGATA	1099	
Db	7693	TTACAGCGGTTCAATGTCGATGAGAAAGACATACACTACCAAGGGGTGTAAGCGAATGCA	7752	
Qy	1100	TTTGTTGGGCTCCGCTGTGCTATTAATCAAGCGCTTGAATCAAGACACTTATAGTGGG	1159	
Db	7753	GGAGACTGATGCGCTCCGATGATATCTTTCAATGTGAAGAGCGTGTGATGATGAAG	7812	
Qy	1160	CTCATGACTCTGCGCAATTTTGGCTGGAACCGTTTCGGGAGGAGCTTCTCTTACAGGA	1219	
Db	7813	CCTCAGAAAGACGGCAGCTTTGGCTCCGAGAGTTGGAATTCATCTTCCATGCGGTT	7872	
Qy	1220	TAGAAGGATTTTCTGGGCAACATCTCTGTGTGT---AAACATGAGTTTCTCTC	1275	
Db	7873	TGTCGAGATCTTCAGCTGTGTATATCGAGTCTGTGGCAAGCATGCAAAACACCC	7932	
Qy	1276	ATGATTTGGGACCTATTTCAAGCTAATGGAAATAGATCTGGGGGTTTGTGTCCAGTTT	1335	
Db	7933	ATGATTTTCGAGGCTTTCCGGAATCTAGAGTTGGGATAGCGCCGAGTTTGTCTC-TATTA	7991	
Qy	1336	GAAAGCGGTTTATTAAGATCCTCGGCTGTGTCATCAACGATAGAAATGAAAGAAATCAGCG	1399	
Db	7992	AACTGTTGTTTTCACAGATCTGTGACTGTGATTAATATGCTACGAGAGAACAGCAT	8051	
Qy	1396	ATGTGCCCTGAAGAAATCTCAGAACTTCCACGTGGATCGCATCTGAAGTGGTTAACGT	1455	
Db	8052	CTTTCTATTTGGGGGTTCAACTTTTGAAGGCTGTAGTCGATGAATATTCAGAAA	8111	
Qy	1456	GTCATCTGAGACCGACGATATGCCATGTTCAAGTCAGGCGGATTCAGAAAGAAAAGCA	1515	
Db	8112	AGCCATGGAAAGCAAGCATCTGTTTATCCCGTCGTGAATAGCCAAAGAGGCGGG	8171	
Qy	1516	AAATTAAGATTAAGGCTTAAGAGCGGGATATCTGAACCTTAATGAATGATGGTGTCACA	1575	
Db	8172	AGATTTGAAGGATGCTTGAACACGGTATTTGCGTGTTTTGAACAGCTATCATTTGGC	8233	
Qy	1576	TCTGACTCGCAAAATATCCAACTCAGGCAATTGCTGACATGCGATGCAATATTTTCAG	1639	
Db	8232	GGCAGTGTGAGGCGCGCTACAGTTGATTAACAGACTGGCGGGGATGAGACCTTCTTCAAC	8291	
Qy	1636	GCACCATGAACCAAGCGGTGATTAACCAACCATATGACAGATGCTCAAACTCTTCTG	1699	

[illegible]

Db 1279 GGCTGCAATACCTTTGCAAAACGGCTGTGACGATTTCTTATCTGAGGACGCGGG 1338  
QY 810 TCGGTGAGGAGCAAGCTTTGTCATGCTTTAGGAGCGCTCTAGTG 860  
Db 1339 CCGGGTGGGCGACGCAATTGTCTCATACCGCTCAGCAATATATGAAAGG 1389

## RESULT 11

ABL15662  
ID ABL15662 standard; cDNA; 4800 BP.

AC ABL15662;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 41468.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-656860/75.

DR P-PSDB; ABB71559.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions -

PS Claim 1; SEQ ID NO 41468; 21bp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 4800 BP; 1265 A; 1116 C; 1141 G; 1278 T; 0 other;

Query Match 2.2%; Score 39.8; DB 23; Length 4800;  
Best Local Similarity 52.0%; Pred. No. 0.2;  
Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 690 TCCGGAGATGTTCTTAAGCCGAAAGTGGCGGTGATTTGGCTGCTTTCCGACTCGT 749

Db 4496 TTCAGGGAATTAATACCAACGTTAAGATCGATCATTTGAGCTGGAATGGCTGCTTATC 4555

QY 750 GGTGGCAACGAAGCTGTTGATGCTGGGGTAGACGATGTATCAATATATGAACAAGTGA 809

Db 4556 GGTGCAAAATCACTTTTGCAAAACGGCTGTGACGATTTCTTATCTGAGGACAGGGG 4615

QY 810 TCGTGTGAGGCAAGCTTTGTACATGCTTTCAAGGACGCTCTAGTGT 860

Db 4616 CCGGGTGGGCGACGATTTGTTCATACCGCTCAGCAATATATCAGAAGGT 4666

## RESULT 12

ABL13667  
ID ABL13667 standard; cDNA; 1872 BP.

AC ABL13667;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35483.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI; 2001-656860/75.

DR P-PSDB; ABB69564.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions -

PS Claim 1; SEQ ID NO 35483; 21bp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1872 BP; 499 A; 477 C; 524 G; 372 T; 0 other;

Query Match 2.2%; Score 39.2; DB 23; Length 1872;  
Best Local Similarity 52.4%; Pred. No. 0.17;  
Matches 86; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 690 TCCGGAGATGTTCTTAAGCCGAAAGTGGCGGTGATTTGGCTGCTTTCCGACTCGT 749

Db 219 TTCAGGGAATTAATACCAACGTTAAGATCGATCATTTGAGCTGGAATGGCTGCTTATC 278

QY 750 GGTGGCAACGAAGCTGTTGATGCTGGGGTAGACGATGTATCAATATATGAACAAGTGA 809

Db 279 GGTGCAAAATCACTTTTGCAAAACGGCTGTGACGATTTCTTATCTGAGGACAGGGG 338

QY 810 TCGTGTGAGGCAAGCTTTGTACATGCTTTCAAGGACGCTC 853

Db 339 CCGGGTGGGCGACGATTTGTTCATACCGCTCAGCAATATATC 382

## RESULT 13

AA15122  
ID AA15122 standard; cDNA to mRNA; 1575 BP.

XX  
AC AAX15122;  
XX  
DT 23-APR-1999 (first entry)  
XX  
DE Nucleic acid encoding an apoptosis inducing protein.  
XX  
KM Cell death; apoptosis; inhibition; proliferation; cancer cell;  
XX apoptosis inducing protein; AIP; chub mackerel; anticancer; ss.  
XX  
OS Scomber japonicus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1575  
FT /\*tag= a  
XX  
PN MO9852972-A1.  
XX  
PD 26-NOV-1998.  
XX  
PF 22-MAY-1998; 98MO-JP02261.  
XX  
PR 23-MAY-1997; 97JP-0133549.  
XX  
PA (TENS-) TENSEI SUISAN CO LTD.  
XX  
PI Iwamoto M, Jung S;  
XX  
DR WPI; 1999-070139/06.  
XX  
DR P-PSDB; AAW96805.  
XX  
PT Apoptosis inducing protein from mackerel and gene encoding it - for  
PT use as anticancer agents  
XX  
PS Claim 5; Page 32-34; 47pp; Japanese.  
XX  
CC The present sequence encodes a protein which induces cell death  
CC (apoptosis) and inhibits the proliferation of cancer cells. The  
CC protein (apoptosis inducing protein, AIP) is isolated from chub  
CC mackerel. The protein can be used as an anticancer agent and as a  
CC reagent for study of the mechanisms of apoptosis in vitro.  
XX  
SQ Sequence 1575 BP; 436 A; 358 C; 406 G; 375 T; 0 other;  
XX  
Query Match 2.1%; Score 38.6; DB 20; Length 1575;  
Best Local Similarity 66.1%; Pred. No. 0.25; Mismatches 34; Indels 3; Gaps 1;  
Matches 72; Conservative 0;  
XX  
QY 713 AAGTGGCGGTCAATGGCGCTGGCATTTCCGAGCTCGTGTGGCAAGCACTGCTTCATG 772  
DB 182 ATGTGTTATAGTCGAGCTGGCATGGCCGAGCTGAGCGCGCCAACTTACTGCAAGACG 241  
XX  
QY 773 CTGGGCTAGACGATGTTACAATATATGAGCAAGTATGCTGTTGAGG 821  
DB 242 CAGG---ACACAGCGTAACCATATGTGAGGCTAATGATCGTGTGAGG 287  
XX  
RESULT 14  
AB256743  
ID AB256743 standard; cDNA; 468 BP.  
XX  
AC AB256743;  
XX  
DT 28-MAR-2003 (first entry)  
XX  
DE Aspergillus oryzae polynucleotide SEQ ID NO 5856.  
XX  
KM Aspergillus oryzae; fermentation; fungus; industrial; EST;  
XX expressed sequence tag; gene; ss.  
XX  
OS Aspergillus oryzae.  
XX  
PN WO200279476-A1.

XX  
PD 10-OCT-2002.  
XX  
XX 22-MAR-2002; 2002MO-IB00890.  
XX  
PR 30-MAR-2001; 2001JP-0098371.  
XX  
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX (NARE-) NAT RES INST BREWING.  
XX (NORQ) NAT FOOD RES INST MIN AGRIC.  
XX  
PI Machida M, Akita O, Kaehiwagi Y, Kitamoto K, Horiuchi H;  
XX Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;  
XX  
DR WPI; 2003-046817/04.  
XX  
PT Detection of expression of specific Aspergillus genes for monitoring  
PT the fermentation and growth conditions of the fungus, using DNA probes  
XX  
PS Claim 1; SEQ ID NO 5856; 48pp + Sequence Listing; Japanese.  
XX  
CC The invention relates to a polynucleotide having any of 6006 specific  
CC sequences (AB250888-AB256893), which are expressed by a fungus under  
CC specific culture conditions including one or more of eutrophic,  
CC oligotrophic, solid, early germination, alkaline, high temperature, low  
CC temperature or maltose culture or polynucleotides stringently hybridising  
CC to these sequences. The polynucleotides are useful for monitoring the  
CC progress of fermentation and the growth conditions of a fungus,  
CC especially of Aspergillus oryzae which is widely used in industrial  
CC fermentation. Also monitoring for fungal contamination.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 468 BP; 124 A; 141 C; 111 G; 90 T; 2 other;  
XX  
Query Match 2.1%; Score 38.4; DB 25; Length 468;  
Best Local Similarity 53.8%; Pred. No. 0.14; Mismatches 82; Indels 3; Gaps 1;  
Matches 100; Conservative 1;  
XX  
QY 706 AAGCCGAAGTGGCGGTCAATGGCGCTGGCATTTCCGAGCTCGTGTGGCAAGCACTG 765  
DB 164 AACACCAATGTCATCGGATCGGCGCGCATCTCAGCCTCACAGCCGCTCTCTC 223  
XX  
QY 766 CTTCATGCTGGGTAGACCATGTTACAATATATGAGCAAGTATGCTGTTGAGGCAAG 825  
DB 224 GATAGCGTGGGATCCCAATTTGGAAATATCTGAAAGCGATACCGGTAAGAGGCGGT 283  
XX  
QY 826 CTTTGGTCAATGCTTTCAAGGAGCGCTCTAGTGTGCGCCGAATGGGGCGATGCGA 885  
DB 284 TTCCGAACCAACTCGTAGGAAACAGC---GCGAATGGGCCAATATGGGCGCCATGCGG 340  
XX  
QY 886 TTTTCT 891  
DB 341 CTGCCT 346  
XX  
RESULT 15  
AAC05199  
ID AAC05199 standard; cDNA; 273 BP.  
XX  
AC AAC05199;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 9274.  
XX  
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 21:01:51 ; Search time 3836.59 Seconds  
(without alignments)  
11409.196 Million cell updates/sec

Title: US-09-434-837-10\_COPY\_1\_1801  
Perfect score: 1801  
Sequence: 1 atgcagcttcacccctccct.....ctatgagtcgagatccga 1801

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estov:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hnc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hnc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pln:*
20:	em_gss_vrt:*
21:	em_gss_fun:*
22:	em_gss_mam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_rod:*
26:	em_gss_phg:*
27:	em_gss_vrl:*
28:	gb_gss1:*
29:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.8	2.8	618	9	AU239500 AU239500
2	44.8	2.5	564	12	BJ123611 BJ123611
3	44.2	2.5	481	12	BJ101485 BJ101485
4	44.2	2.5	486	12	BJ103567 BJ103567

5	44.2	2.5	508	9	AU199996 AU199996
6	44.2	2.5	511	12	BJ108542 BJ108542
7	44.2	2.5	513	9	AU203802 AU203802
8	44.2	2.5	529	12	BJ120847 BJ120847
9	44.2	2.5	540	9	AU201678 AU201678
10	44.2	2.5	560	14	CB401025 CB401025
11	44.2	2.5	584	12	BJ121613 BJ121613
12	44.2	2.5	586	12	BJ121635 BJ121635
13	44.2	2.5	597	9	AU207635 AU207635
14	44.2	2.5	603	12	BJ101145 BJ101145
15	44.2	2.5	608	9	AU204573 AU204573
16	44.2	2.5	619	12	BJ117235 BJ117235
17	44.2	2.5	634	12	BJ124022 BJ124022
18	44.2	2.5	655	12	BJ103250 BJ103250
19	43.8	2.4	1201	13	BX381961 BX381961
20	43.4	2.4	482	12	BJ101110 BJ101110
21	43	2.4	338	9	AU109666 AU109666
22	42.6	2.4	631	12	BJ117453 BJ117453
23	42.4	2.4	860	13	BU348455 BU348455
24	42	2.3	362	14	CA926583 CA926583
25	42	2.3	1201	13	BX376097 BX376097
26	40.8	2.3	999	29	CNS060UH CNS060UH
27	40.6	2.3	310	13	BU821300 BU821300
28	40.6	2.3	411	9	A1102326 A1102326
29	40.6	2.3	479	14	CA934750 CA934750
30	40.6	2.3	1376	29	AG176243 AG176243
31	39.4	2.2	284	9	A1101650 A1101650
32	39.4	2.2	466	9	AA943462 AA943462
33	39.4	2.2	462	12	BI300778 BI300778
34	39.4	2.2	511	10	BG378762 BG378762
35	39.4	2.2	560	12	BI296551 BI296551
36	39.2	2.2	404	12	BJ105833 BJ105833
37	39.2	2.2	410	12	BI611075 BI611075
38	39.2	2.2	437	13	BY467572 BY467572
39	39.2	2.2	501	14	CA614930 CA614930
40	39.2	2.2	502	12	BI368223 BI368223
41	39.2	2.2	554	9	A1457056 A1457056
42	39.2	2.2	564	12	BI566474 BI566474
43	39.2	2.2	622	12	BI627861 BI627861
44	39.2	2.2	623	12	BI588865 BI588865
45	39.2	2.2	657	9	AA440930 AA440930

## ALIGNMENTS

RESULT 1  
LOCUS AU239500  
DEFINITION AU239500 RAFL19 Arabidopsis thaliana CDNA clone RAFL19-89-D06 5', mRNA sequence.  
ACCESSION AU239500  
VERSION AU239500.1 GI:19878669  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE  
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Saitou,M., Nakajima,M., Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Aizawa,T., Shibata,K., Shinozaki,K., M., Hayashizaki,Y. and Shinozaki,K.  
Large scale analysis of Arabidopsis full-length cDNA  
Unpublished  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: msekic@rcc.riken.go.jp



Db 384 ATTCATGCCATACCGTACAGGATG 408

RESULT 4  
LOCUS BJI03567 486 bp mRNA linear EST 18-JAN-2002  
DEFINITION BJI03567 unpublished oligo-capped cDNA library, C. elegans L1 stage  
ACCESSION BJI03567  
VERSION BJI03567.1 GI:18246237  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 486)  
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.  
and Sugano, S.  
A complementary view of the C. elegans genome  
Unpublished  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1..486  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1051f01"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"

BASE COUNT 137 a 104 c 114 g 130 t 1 others

ORIGIN

Query Match 2.5%; Score 44.2; DB 12; Length 486;  
Best Local Similarity 56.6%; Pred. No. 0.28; Mismatches 63; Indels 0; Gaps 0;  
Matches 82; Conservative 0;

LOCUS BJI08542 511 bp mRNA linear EST 30-MAY-2003  
DEFINITION BJI08542 unpublished oligo-capped cDNA library, C. elegans L1 stage  
ACCESSION BJI08542  
VERSION BJI08542.2 GI:31243973  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 511)  
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.  
and Sugano, S.  
A complementary view of the C. elegans genome  
Unpublished  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1..511  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"

BASE COUNT 148 a 109 c 116 g 135 t

ORIGIN

Query Match 2.5%; Score 44.2; DB 9; Length 508;  
Best Local Similarity 56.6%; Pred. No. 0.28; Mismatches 63; Indels 0; Gaps 0;  
Matches 82; Conservative 0;

LOCUS BJI08542 508 bp mRNA linear EST 17-JUL-2001  
DEFINITION BJI08542 unpublished oligo-capped cDNA library, stage L2  
ACCESSION BJI08542  
VERSION BJI08542.1 GI:14827576  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 508)  
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.

and Sugano, S.  
A complementary view of the C. elegans genome  
Unpublished  
Contact: Yuji Kohara  
Genome Biology Lab.  
National Institute of Genetics  
Yata 111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykohara@lab.nig.ac.jp.  
Location/Qualifiers  
1..508  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk774a07"  
/sex="Hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L2"  
/clone\_lib="unpublished oligo-capped cDNA library, stage  
L2"

BASE COUNT 148 a 109 c 116 g 135 t

ORIGIN

Query Match 2.5%; Score 44.2; DB 9; Length 508;  
Best Local Similarity 56.6%; Pred. No. 0.28; Mismatches 63; Indels 0; Gaps 0;  
Matches 82; Conservative 0;

LOCUS BJI08542 511 bp mRNA linear EST 30-MAY-2003  
DEFINITION BJI08542 unpublished oligo-capped cDNA library, C. elegans L1 stage  
ACCESSION BJI08542  
VERSION BJI08542.2 GI:31243973  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 511)  
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.  
and Sugano, S.  
A complementary view of the C. elegans genome  
Unpublished  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1..511  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"

BASE COUNT 148 a 109 c 116 g 135 t

ORIGIN

Query Match 2.5%; Score 44.2; DB 9; Length 508;  
Best Local Similarity 56.6%; Pred. No. 0.28; Mismatches 63; Indels 0; Gaps 0;  
Matches 82; Conservative 0;

LOCUS BJI08542 508 bp mRNA linear EST 17-JUL-2001  
DEFINITION BJI08542 unpublished oligo-capped cDNA library, stage L2  
ACCESSION BJI08542  
VERSION BJI08542.1 GI:14827576  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 508)  
Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.

```

/clone="yk1109b04"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT      150 a      108 c      117 g      135 t      1 others
ORIGIN

Query Match      2.5%; Score 44.2; DB 12; Length 511;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 706 AAGCCGAAGTGGGGTCATTTGGCGGATTCCTCCGACTCGTGTGGCAAGAACTG 765
Db 209 AAACCGTCTATAGCAATGTCGGCGCGGTATTTCCGACTATCGACCGCGCTGCACTG 268
Qy 766 CTTCATCTCGGGGTAGACGATGTTACATATATAGCAAGTATCGTGTGGAGCGAAG 825
Db 269 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGCGACA 328
Qy 826 CTTTGTGCATCTCTTCAGGACG 850
Db 329 ATTATGCCATACCGTACAGGATG 353

RESULT 7
AU203802
LOCUS AU203802 unpublished oligo-capped cDNA library, stage L1
DEFINITION Caenorhabditis elegans cDNA clone yk825e11 5', mRNA sequence.
ACCESSION AU203802
VERSION AU203802.2 GI:31231885
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
1 (bases 1 to 513)
Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C.elegans genome
Unpublished
On Jul 17, 2001 this sequence version replaced gi:14834726.
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
source
1..513
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk825e11"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, stage
L1"
BASE COUNT      141 a      114 c      114 g      143 t      1 others
ORIGIN

Query Match      2.5%; Score 44.2; DB 9; Length 513;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 706 AAGCCGAAGTGGGGTCATTTGGCGGATTCCTCCGACTCGTGTGGCAAGAACTG 765
Db 230 AAACCGTCTATAGCAATGTCGGCGCGGTATTTCCGACTATCGACCGCGCTGCACTG 269

```

```

Qy 766 CTTCATCTCGGGGTAGACGATGTTACATATATAGCAAGTATCGTGTGGAGCGAAG 825
Db 290 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGCGACA 349
Qy 826 CTTTGTGCATCTCTTCAGGACG 850
Db 350 ATTATGCCATACCGTACAGGATG 374

RESULT 8
BU120847
LOCUS BU120847 529 bp mRNA linear EST 23-JAN-2002
DEFINITION BU120847 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone yk1261a07 5', mRNA sequence.
ACCESSION BU120847
VERSION BU120847.1 GI:18280980
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
1 (bases 1 to 529)
Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
and Sugano, S.
A complementary view of the C.elegans genome
Unpublished
Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
source
1..529
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1261a07"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT      154 a      112 c      121 g      142 t

Query Match      2.5%; Score 44.2; DB 12; Length 529;
Best Local Similarity 56.6%; Pred. No. 0.28;
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 706 AAGCCGAAGTGGGGTCATTTGGCGGATTCCTCCGACTCGTGTGGCAAGAACTG 765
Db 193 AAACCGTCTATAGCAATGTCGGCGCGGTATTTCCGACTATCGACCGCGCTGCACTG 252
Qy 766 CTTCATCTCGGGGTAGACGATGTTACATATATAGCAAGTATCGTGTGGAGCGAAG 825
Db 253 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGCGACA 312
Qy 826 CTTTGTGCATCTCTTCAGGACG 850
Db 313 ATTATGCCATACCGTACAGGATG 337

RESULT 9
AU201678
LOCUS AU201678 540 bp mRNA linear EST 17-JUL-2001
DEFINITION AU201678 unpublished oligo-capped cDNA library, stage L1
Caenorhabditis elegans cDNA clone yk797d07 5', mRNA sequence.
ACCESSION AU201678
VERSION AU201678.1 GI:14830776

```

KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
REFERENCE 1 (bases 1 to 540)  
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished  
COMMENT Contact: Yuji Kohara  
Genome Biology Lab.  
National Institute of Genetics  
Yata 1111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykohara@lab.nig.ac.jp.  
Location/Qualifiers  
FEATURES  
source  
1. 540  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="YK797d07"  
/sex="Hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
/clone\_1lb="unpublished oligo-capped cdna library, stage L1"  
BASE COUNT 157 a 117 c 119 g 147 t  
ORIGIN  
Query Match 2.5%; Score 44.2; DB 9; Length 540;  
Best Local Similarity 56.6%; Pred. No. 0.28;  
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 706 AACCCGAAGTGGCGGTGCGTGCATTTGCGACTTCGATGCGCAACGAACTG 765  
|||  
DB 229 AACCGTCTATGACATCGTCGCGCGGATTTCCGACTATCGACGCGCGTCACTG 288  
|||  
QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATGCTGTGGAGGCAAG 825  
|||  
DB 289 ATGAACCTTGGAATGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGGCAAG 348  
|||  
QY 826 CTTTGTCAATGCTTTTCAGGAGC 850  
|||  
DB 349 ATTCAATGCCATACCGTACAGAGATG 373  
|||  
RESULT 10  
CB401025 560 bp mRNA linear EST 15-MAY-2003  
LOCUS OSTF186F11\_1 AD-wtmcDNA Caenorhabditis elegans cDNA, mRNA sequence.  
DEFINITION CB401025  
ACCESSION CB401025  
VERSION CB401025.1 GI:30742752  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
REFERENCE 1 (bases 1 to 560)  
AUTHORS Reboul,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M., Armstrong  
,C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T., Hudson  
,J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,  
Endress,G.A., Jena,S., Chevret,E., Papasotiriopoulos,V., Tolias,P.P.,  
Plescek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,  
Doucette-Stamm,L., Hill,D.E. and Vidal,M.  
TITLE C. elegans ORFome version 1.1: experimental verification of the  
genome annotation and resource for proteome-scale protein  
expression  
JOURNAL Nat. Genet., (2003) In press  
COMMENT Contact: Vidal M  
Marc Vidal laboratory

Dana Farber Cancer Institute  
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
Tel: 617 632 5180  
Fax: 617 632 5739  
Email: Marc.Vidal@dfci.harvard.edu  
Sequence tag of Gateway entry clones. The primers used were  
designed on the predicted protein encoding ORF. C. elegans ORFome  
cloning project : Contact david\_hill@dfci.harvard.edu or  
marc.vidal@dfci.harvard.edu  
POLYA=No.  
FEATURES  
source  
1. 560  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/sex="Hermaphrodite and male"  
/tissue\_type="whole animal"  
/dev\_stage="mixed stage"  
/clone\_1lb="AD-wtmcDNA"  
/note="The AD-wtmcDNA library was generated with poly(A)+  
RNA isolated from both hermaphrodite and male N2 worms of  
all larval stages, embryos, adults and dauers and the  
subsequent generation of cDNAs by poly(A) priming. The  
cDNAs were cloned into pC086"  
BASE COUNT 175 a 116 c 122 g 147 t  
ORIGIN  
Query Match 2.5%; Score 44.2; DB 14; Length 560;  
Best Local Similarity 56.6%; Pred. No. 0.28;  
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 706 AACCCGAAGTGGCGGTGCGTGCATTTGCGACTGCTGTGCGCAACGAACTG 765  
|||  
DB 69 AACCGTCTATGACATCGTCGCGCGGATTTTCGACTATCGACCGCGTCACTG 128  
|||  
QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATGCTGTGGAGGCAAG 825  
|||  
DB 129 ATGAACCTTGGAATGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGGCAAG 188  
|||  
QY 826 CTTTGTCAATGCTTTTCAGGAGC 850  
|||  
DB 189 ATTCAATGCCATACCGTACAGAGATG 213  
|||  
RESULT 11  
BJ121613 584 bp mRNA linear EST 23-JAN-2002  
LOCUS BJ121613  
DEFINITION BJ121613 unpublished oligo-capped cDNA library, C. elegans L1 stage  
Caenorhabditis elegans cDNA clone yk1278101 5', mRNA sequence.  
ACCESSION BJ121613  
VERSION BJ121613.1 GI:18281750  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
REFERENCE 1 (bases 1 to 584)  
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished  
COMMENT Contact: Tadao Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6855  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp.  
Location/Qualifiers  
FEATURES  
source  
1. 584  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"

Query Match	2.5%;	Score 44.2;	DB 12;	Length 584;	
Best Local Similarity	56.6%;	Pred. No. 0.29;			
Matches	82;	Conservative	0;	Mismatches	63;
				Indels	0;
				Gaps	0;
Db	210	AAACCGTATATAGCAATGCTCGCGCCGGATTTTCCGGACTATGACCGCGCTGCAC	269		
Qy	706	AAGCCGAAGTGGCGGCTGATTTGGCGCTTGCGACTTTCCGACTCGTGTGGCAAGCAACTG	765		
Db	270	ATTGAACTTGAATTTGACGATTTTGATATCTACAGAGGCTCTGACCGGATGGAGGACGA	329		
Qy	826	CTTGTGTCACATGCTTTGAGGAGC	850		
Db	330	ATTATGCCATACCGTACAGAGATG	354		
RESULT 12					
Bj121635					
LOCUS	Bj121635	586 bp	mRNA	linear	EST 23-JAN-2002
DEFINITION	Bj121635	unpublished oligo-capped cDNA library, C. elegans L1 stage			
VERSION	Bj121635	Caenorhabditis elegans cDNA clone yk1278h02.5, mRNA sequence.			
KEYWORDS	Bj121635.1	GI:18281772			
SOURCE	EST.				
ORGANISM	Caenorhabditis elegans				
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea				
	; Rhabditidae; Peloderae; Caenorhabditis.				
REFERENCE	1.	(bases 1 to 586)			
AUTHORS	Kohara, Y., Shih, I.-T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.				
	and Sugano, S.				
	A complementary view of the C. elegans genome				
TITLE	Unpublished				
JOURNAL	Contact: Tadasu Shin-i				
COMMENT	Center For Genetic Resource Information				
	National Institute of Genetics				
	111 Yata, Mishima, Shizuoka 411-8540, Japan				
	Tel: 81-559-81-6856				
	Fax: 81-559-81-6855				
	Email: tshini@genes.nig.ac.jp.				
	Location/Qualifiers				
FEATURES	1..586				
source	/organism="Caenorhabditis elegans"				
	/mol_type="mRNA"				
	/strain="N2"				
	/db_xref="taxon:6239"				
	/clone="yk1278h02"				
	/sex="hermaphrodite"				
	/tissue_type="whole animal"				
	/dev_stage="L1"				
	/clone_lib="unpublished oligo-capped cDNA library, C.				
	elegans L1 stage"				
BASE COUNT	179 a	123 c	130 g	154 t	
ORIGIN					
Query Match	2.5%;	Score 44.2;	DB 12;	Length 586;	
Best Local Similarity	56.6%;	Pred. No. 0.29;			
Matches	82;	Conservative	0;	Mismatches	63;
				Indels	0;
				Gaps	0;
Qy	706	AAGCCGAAGTGGCGGCTGATTTGGCGCTTGCGACTTTCCGGACTATGACCGCGCTGCAC	765		

Db	210	AAACCCGCTTAGCAATCGTCGGCGCGGCGATTTCCGGACTATCGACCGCGCGTCGACTG	265
Qy	766	CTTCATCTCTGGGGAGACGATGTTACAATATATGAAGCAAGTATCGTGTGGAGCGAAG	825
Db	270	ATTGAACCTTGAAATGACGATTTGATATCTACGAAGGTCTCGACCGATCGAGGCGACA	329
Qy	826	CTTTGGTCACATGCTTTCAGGAGCG	850
Db	330	ATTGATGCCATACCGTACAAAGATG	354
RESULT 13			
LOCUS	AU207635	597 bp	mRNA linear EST 17-JUN-2001
DEFINITION	AU207635	unpublished oligo-capped cDNA library, stage L1	
LOCUS	AU207635	Caenorhabditis elegans cDNA clone YK1004b11 5', mRNA sequence.	
ACCESSION	AU207635		
VERSION	AU207635.1	GI:14841711	
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae		
AUTHORS	1 (bases 1 to 597)		
	Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.		
	and Sugano, S.		
TITLE	A complementary view of the C.elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara		
	National Institute of Genetics		
	Yata 111, Mishima, Shizuoka 411, Japan		
	Tel: 81-559-81-6854		
	Fax: 81-559-81-6855		
	Email: ykohara@lab.nig.ac.jp.		
	Location/Qualifiers		
FEATURES	1..597		
source	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="YK1004b11"		
	/sex="Hermaphrodite"		
	/tissue_type="whole animal"		
	/dev_stage="L1"		
	/clone_lib="unpublished oligo-capped cDNA library, stage L1"		
BASE COUNT	179 a 126 c 131 g 161 t		
ORIGIN			
Query Match	2.5%; Score 44.2; DB 9; Length 597;		
Best Local Similarity	56.6%; Pred. No. 0.29; 63; Indels 0; Gaps 0;		
Matches	82; Conservative 0; Mismatches		
Qy	706	AAGCGGAAGGGGGGCGATTCATTCGGGCGGCGACTTCGGAGCTCGTGTGGCAACGAAGT	765
Db	223	AAACCGCTTAGCAATCGTCGGCGCGGCGATTTCCGACATACGACCGCGCTCGACTG	282
Qy	766	CTTCATCTCTGGGGAGACGATGTTACAATATATGAAGCAAGTATCGTGTGGAGCGAAG	825
Db	283	ATTGAACCTTGAAATGACGATTTGATATCTACGAAGGTCTCGACCGATCGAGGCGACA	342
Qy	826	CTTTGGTCACATGCTTTCAGGAGCG	850
Db	343	ATTGATGCCATACCGTACAAAGATG	367
RESULT 14			
LOCUS	BJ101145	603 bp	mRNA linear EST 18-JAN-2002
DEFINITION	BJ101145	unpublished oligo-capped cDNA library, C. elegans L1 stage	
ACCESSION	BJ101145	Caenorhabditis elegans cDNA clone YK1022g05 5', mRNA sequence.	

VERSION	KEYWORDS	EST.	Bj101145.1	GI:18243815
SOURCE	ORGANISM	Caenorhabditis elegans	Caenorhabditis elegans	
REFERENCE	AUTHORS	Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.		
TITLE	JOURNAL	Unpublished		
COMMENT		Contact: Tadasu Shin-i National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp. Location/Qualifiers		
FEATURES	Source	1. 603 /organism="Caenorhabditis elegans" /mol_type="mRNA" /strain="N2" /db_xref="taxon:6239" /clone="yk1022g05" /sex="hermaphrodite" /tissue_type="whole animal" /dev_stage="L1" /clone_lib="unpublished oligo-capped cdna library, C. elegans L1 stage"		
BASE COUNT	Origin	180 a 130 c 131 g 162 t		
Query Match		2.5%; Score 44.2; DB 12; Length 603; Best Local Similarity 56.6%; Pred. No. 0.29;		
Matches		82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;		
QY		706 AAGCGAAAGTGGCGGCTGATTCGGCGCTGGCAATTCGCGACTCGTGTGGCAACGAAGT 765		
Db		229 AAACCGTTATATAGCAATGTCGGCGCCCGTATTTCCGAGTATGACACGCGCGTGCAGTG 288		
QY		766 CTTATGTCGGGGTATAGCATGTTATATATATATAGCAAGTATCGTGTGGAGGCAAG 825		
Db		289 ATTGAACCTTGAAATGACGATTTTGATATCTTACGAAGGCTCTGACCGGATGGAGGAGA 348		
QY		826 CTTTGTCACATGCTTTCAGGAGC 850		
Db		349 ATTATGCGATACCGTACAGAGATG 373		
RESULT 15				
LOCUS		AU204573 608 bp mRNA linear EST 17-JUL-2001		
DEFINITION		AU204573 unpublished oligo-capped cDNA library, stage L4		
ACCESSION		Caenorhabditis elegans cDNA clone yk839b11 5', mRNA sequence.		
VERSION		AU204573		
KEYWORDS		AU204573.1 GI:14836084		
SOURCE		EST.		
ORGANISM		Caenorhabditis elegans		
		Caenorhabditis elegans		
		Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea		
		; Rhabditidae; Peloderinae; Caenorhabditis.		
REFERENCE		1 (bases 1 to 603)		
AUTHORS		Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.		
TITLE		A complementary view of the C.elegans genome		
JOURNAL		Unpublished		
COMMENT		Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855		

Email: ykohana@lab.nig.ac.jp.  
Location/Qualifiers

source

1. 608  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="YK839B11"  
/sex="Hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L4"  
/clone\_idb="unpublished oligo-capped cDNA library, stage L4"

BASE COUNT      188 a    129 c    131 g    160 t

ORIGIN

Query Match                  2.5%: Score 44.2; DB 9; Length 608;  
Best Local Similarity 56.6%; Pred. No. 0.29; Mismatches 63; Indels 0; Gaps 0;  
Matches 82; Conservative 0;

Dy AAGCCGAAGGCGGCATCTATTGGCGCTGGCACTTCGGACTCGTGTGGCAACGAACTGTG 765  
|||  
Db 173 AAACCGCTATAGCAATCGTCGGCGGCCGGATTTCCGACATATGACC GCCCGCTCAGCTG 232  
|||  
Dy 766 CTTGATCTGGGGTAGACGATGTTACAATATATGAACCAACTGATCGTGTTGAGGCAAG 825  
|||  
Db 233 ATTGAACCTTGAATATGACGATTTTGATATATCAAGAAGTCTCGACCGGATCGAGGACGA 292  
|||  
Dy 826 CTTTGTCACATGCTTCAGGAGAG 850  
|||  
Db 293 ATTCATGCATACCGTACAAAGGATG 317

Search completed: November 24, 2003, 04:01:40  
Job time : 3837.59 secs

```

Query Match      2.5%: Score 44.2; DB 9; Length 608;
Best Local Similarity 56.6%; Pred. No. 0.29; Mismatches 0; Gaps 0.
Matches 82; Conservative 0; Indels 0;

OY      706 AAGCCGAAAGTGGCGGCTCATTTGGCGCTGGCAATTTCCGGACTCGTGTCGAACGAACTG 765
          |||||
DB       173 AAACCGTCTATAGCAATCGTCGGCGCGGTATTTCCGGACTATCGACCGCGCTGCACTG 232
          |||||

OY      766 CTTATGCTGGGGTAGACGATGTTACAATATATGAAACAAGTGAATCGTGTGGAGGCCAAG 825
          |||||
DB       233 ATTGAACCTTGGAAATTGACGATTTTGATATATTAAGAAGGTCTCGACCGGATCGAGGCAGA 292
          |||||

OY      826 CTTTGTTCACATGCTTTCAGGAGC 850
          |||||
DB       293 ATTCATGCCATACCCTACAAAGATG 317
          |||||

Search completed: November 24, 2003, 04:01:40
Job time : 3837.59 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 20:32:25 ; Search time 105.342 Seconds  
(without alignments)  
7546.168 Million cell updates/sec

Title: US-09-434-837-10\_COPY\_1\_1801

Perfect score: 1801

Sequence: 1 atgtcagcttcacacctctcct.....ctatgagtcgacgaccca 1801

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTCUTS COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1752.4	97.3	24595	6	5428147-1
2	1697.4	94.2	2211	3	US-09-073-587-2
3	121.6	6.8	2279	1	US-08-008-216-8
4	121.6	6.8	2279	1	US-08-459-569-8
5	121.6	6.8	2279	1	US-08-458-831-8
6	121.6	6.8	21126	1	US-08-008-216-19
7	121.6	6.8	21126	1	US-08-459-569-19
8	121.6	6.8	21126	1	US-08-458-831-19
9	38.6	2.1	1575	3	US-09-230-388-2
10	37	2.1	4403765	3	US-09-103-840A-2
11	37	2.1	4411529	3	US-09-103-840A-1
12	35.2	2.0	1032	4	US-09-415-277C-4
13	35.2	2.0	1032	4	US-09-415-277C-6
14	35.2	2.0	3593	4	US-09-404-627-3
15	35.2	2.0	4205	4	US-09-404-627-1
16	35	1.9	7218	1	US-08-232-463-14
17	34.6	1.9	615	4	US-09-252-991A-8485
18	34.6	1.9	1056	4	US-09-252-991A-8526
19	34.6	1.9	1830	4	US-09-252-991A-8526
20	33.6	1.9	1617	4	US-09-252-991A-8322
21	33.6	1.9	1665	4	US-09-252-991A-6497
22	33.6	1.9	419	1	US-08-486-013-59
23	33.4	1.9	419	2	US-08-482-279-59
24	33.4	1.9	419	2	US-08-342-268-59
25	33.4	1.9	419	3	US-09-015-968-59
26	33.4	1.9	419	4	US-09-397-386-59
27	33.4	1.9	7218	1	US-08-232-463-14

28	33.4	1.9	536165	4	US-09-214-808-1	Sequence 1, Appl1
29	33.2	1.8	1883	4	US-09-063-733A-45	Sequence 45, Appl1
30	33.2	1.8	1950	4	US-09-063-733A-50	Sequence 50, Appl1
31	33.2	1.8	2700	4	US-09-063-733A-15	Sequence 15, Appl1
32	32.4	1.8	53332	4	US-09-801-861-3	Sequence 3, Appl1
33	32.4	1.8	87563	4	US-09-453-702B-57	Sequence 57, Appl1
34	32	1.8	754	2	US-08-879-995A-2	Sequence 2, Appl1
35	32	1.8	754	2	US-09-215-096-2	Sequence 2, Appl1
36	32	1.8	972	4	US-09-018-584A-40	Sequence 358, App
37	32	1.8	1000	3	US-08-341-916-1	Sequence 40, Appl1
38	32	1.8	1482	1	US-08-805-166-1	Sequence 1, Appl1
39	32	1.8	1482	1	US-08-805-169-1	Sequence 1, Appl1
40	32	1.8	1482	1	US-08-805-169-1	Sequence 1, Appl1
41	32	1.8	1482	2	US-08-957-365-1	Sequence 1, Appl1
42	32	1.8	1482	2	US-08-957-365-1	Sequence 1, Appl1
43	32	1.8	1602	3	US-08-476-123-11	Sequence 11, Appl1
44	32	1.8	1605	3	US-09-124-541-3	Sequence 3, Appl1
45	31.6	1.8	2304	1	US-08-464-266-1	Sequence 1, Appl1

#### ALIGNMENTS

```
RESULT 1
5428147-1
PATENT NO. 5428147
APPLICANT: BARKER, RICHARD F.; KEMP, JOHN D.
TITLE OF INVENTION: OCTOPINE T-DNA PROMOTERS
NUMBER OF SEQUENCES: 17
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/91,538
  FILING DATE: 13-JUL-1993
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 869,216
  FILING DATE: 13-APR-1992
  APPLICATION NUMBER: 869,216
  FILING DATE: 13-APR-1992
  APPLICATION NUMBER: 440,432
  FILING DATE: 21-NOV-1989
  APPLICATION NUMBER: 553,786
  FILING DATE: 19-NOV-1983
  APPLICATION NUMBER: 741,034
  FILING DATE: 06-AUG-1991
  APPLICATION NUMBER: 144,775
  FILING DATE: 20-JAN-1988
  APPLICATION NUMBER: 485,614
  FILING DATE: 15-APR-1983
  APPLICATION NUMBER: 713,624
  FILING DATE: 10-JUN-1991
  APPLICATION NUMBER: 260,574
  FILING DATE: 21-OCT-1988
  APPLICATION NUMBER: 848,733
  FILING DATE: 01-APR-1986
  APPLICATION NUMBER: 535,354
  FILING DATE: 26-SEP-1983
; SEQ ID NO:1
; LENGTH: 24595
5428147-1
Query Match          97.3%; Score 1752.4; DB 6; Length 24595;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1784; Conservative 0; Mismatches 16; Indels 2; Gaps 2;
```

QY 1 ATGTGAGCTTCACTCTCTTGATTAACAGTGCGATCATCTCCCAACCAAAATGGTGAT 60

DB 5809 ATGTGAGCTTCACTCTCTTGATTAACAGTGCGATCATCTCCCAACCAAAATGGTGAT 5868

QY 61 CTGACATGCTTCACTCTCTTGATTAACAGTGCGATCATCTCCCAACCAAAATGGTGAT 120

DB 5869 CTGACATGCTTCACTCTCTTGATTAACAGTGCGATCATCTCCCAACCAAAATGGTGAT 5928

QY 121 CGAAGAGCTTCTTGGGGAAGAGAGATTACTCAAGCTCCACGAGTCCAGCGCTGGGTTA 180

Db 5929 CGAAGAGCTTCTAGGGAAGAGGATTACTCAATCTCAACGAGTGCAGCGCTGGGTTA 5988  
Qy 181 GCTTGGCAAAAGGCTGGCCGATGGTTCGCTTCCCGAGATCTCAGCTGGTGGAAAGTATGCA 240  
Db 5989 GCTTGGCAAAAGGCTGGCCGATGGTTCGCTTCCCGAGATCTCAGCTGGTGGAAAGTATGCA 6048  
Qy 241 GTTCTCTCCGCTTATATCTATTTATTTGGCAAAATTTGGGGCGGATCTTGATCGAAA 300  
Db 6049 GTTCTCTCCGCTTATATCTATTTATTTGGCAAAATTTGGGGCGGATCTTGATCGAAA 6108  
Qy 301 CCTTGGCGCGGGCAACAGTGAAGTGTCTCGTTCATCGACTTGAACACATTTGCATG 360  
Db 6109 CCTTGGCGCGGGCAACAGTGAAGTGTCTCGTTCATCGACTTGAACACATTTTGATG 6168  
Qy 361 GATTTCTCCGAACACACTAATTCAGACCCCTGTTTTGCTGAGCGGTAAAGATGCGA 420  
Db 6169 GATTTCTCCGAACACACTAATTCAGACCCCTGTTTTGCTGAGCGGTAAAGATGCGA 6228  
Qy 421 CCGATTGATCTTATGCTATTTCTGAGCATTCTCAATCTCAAGACTGCGGCTTTGCAAC 480  
Db 6229 CCGATTGATCTTATGCTATTTCTGAGCATTCTCAATCTCAAGACTGCGGCTTTGCAAC 6288  
Qy 481 CTGCAATGCGCGTGTATCGAAGATGGCAAGATGAATGCGTTACCGGTTTACCATACC 540  
Db 6289 CTGCAATGCGCGTGTATCGAAGATGGCAAGATGAATGCGTTACCGGTTTACCATACC 6348  
Qy 541 CTTGAAGGGGCGGCTGCTATTTGACATGATGCTTATAGTGGAAACCTGATGCTGAAGGT 600  
Db 6349 CTTGAAGGGGCGGCTGCTATTTGACATGATGCTTATAGTGGAAACCTGATGCTGAAGGT 6408  
Qy 601 TCGGCAAGGTTCTTTCCCAACAATCGACTTCTCAACACTGACAGACCGTTTTTTGACCA 660  
Db 6409 TCGGCAAGGTTCTTTCCCAACAATCGACTTCTCAACACTGACAGACCGTTTTTTGACCA 6468  
Qy 661 TGTTCGATATGAGCGAGATCGGCTTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720  
Db 6469 TGTTCGATATGAGCGAGATCGGCTTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 6528  
Qy 721 GTTATTTGGCGCTGGCATTTCCGAGCTGTGTGGTGGCAACGACTGCTTACGCTGGGGTA 780  
Db 6529 GTTATTTGGCGCTGGCATTTCCGAGCTGTGTGGTGGCAACGACTGCTTACGCTGGGGTA 6588  
Qy 781 GACGATGTTCAATATATGAAAGCAAGTATGATGATGTTGAGGAGCAAGCTTTGCTCAATGCT 840  
Db 6589 GACGATGTTCAATATATGAAAGCAAGTATGATGATGTTGAGGAGCAAGCTTTGCTCAATGCT 6648  
Qy 841 TTGAGGAGCGCTCTAGTGTCTGTGGCGAAATGGGGCGATGGCAATTTCTCTGCTGCA 900  
Db 6649 TTGAGGAGCGCTCTAGTGTCTGTGGCGAAATGGGGCGATGGCAATTTCTCTGCTGCA 6708  
Qy 901 TTTCTGCTGTTTTTTCTTCTCGAGCGTTACGGCTGTCTTGATGAGGCGGTTCCCAAT 960  
Db 6709 TTTCTGCTGTTTTTTCTTCTCGAGCGTTACGGCTGTCTTGATGAGGCGGTTCCCAAT 6768  
Qy 961 CCGGAGCAGTGCACACTTACTTGTCTCAACAGGCGTCCCAATCAATGATGAAAGCGGG 1020  
Db 6769 CCGGAGCAGTGCACACTTACTTGTCTCAACAGGCGTCCCAATCAATGATGAAAGCGGG 6828  
Qy 1021 CAGCTGCACCGAAGCTGTTCATCGCGTTTACACGCGTGGCGGTTCTTGAAGAC 1080  
Db 6829 CAGCTGCACCGAAGCTGTTCATCGCGTTTACACGCGTGGCGGTTCTTGAAGAC 6888  
Qy 1081 GGTTCATATAGGAGATATTTGTTGGCTTCGCTGTGCTTATTTACTCAGGCGCTTGA 1140  
Db 6889 GGTTCATATAGGAGATATTTGTTGGCTTCGCTGTGCTTATTTACTCAGGCGCTTGA 6948  
Qy 1141 TCAGGACACTTATGTTGGGCTCATGATCTCTGGCAATTTGGCGAATCGTTTGGGAGG 1200  
Db 6949 TCAGGACACTTATGTTGGGCTCATGATCTCTGGCAATTTGGCGAATCGTTTGGGAGG 7008  
Qy 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTGCTGTGT -AA 1259  
Db 7009 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTGCTGTGTGTGA 7068

Qy 1260 ACATGAGTTTCTCATGATTTGGGACCTATTCAGCTATATGGGAATAGGATCTGGCGG 1319  
Db 7069 ACATGAGTTTCTCATGATTTGGGACCTATTCAGCTATATGGGAATAGGATCTGGCGG 7127  
Qy 1320 GTTGTGCTCAGTTTGTGAAGCGGGTTTATTTAGATCTCTCGCTGGTGTCAACGATA 1379  
Db 7128 GTTGTGCTCAGTTTGTGAAGCGGGTTTATTTAGATCTCTCGCTGGTGTCAACGATA 7187  
Qy 1380 TGAAGAAATACCGGATGTGCTCTGAGGAGATCTCAGAACTTCCAGCTGTGATCGCATC 1439  
Db 7188 TGAAGAAATACCGGATGTGCTCTGAGGAGATCTCAGAACTTCCAGCTGTGATCGCATC 7247  
Qy 1440 TGAAGGTTAAGGTTGTCTGTGAGCCAGCGCATATGCTCATGTTCAAGTCCAGGCGAT 1499  
Db 7248 TGAAGGTTAAGGTTGTCTGTGAGCCAGCGCATATGCTCATGTTCAAGTCCAGGCGAT 7307  
Qy 1500 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTTATGA 1559  
Db 7308 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTTATGA 7367  
Qy 1560 TAAAGTGTGTCAATCTGACTCCGAAATATCCAACTCAGCATTTGCTGACATGCGA 1619  
Db 7368 TAAAGTGTGTCAATCTGACTCCGAAATATCCAACTCAGCATTTGCTGACATGCGA 7427  
Qy 1620 TACCAATATTTTCAAGGACACAGTGAACCAAGGGTTGATTAAGCCATATGACAGGATC 1679  
Db 7428 TACCAATATTTTCAAGGACACAGTGAACCAAGGGTTGATTAAGCCATATGACAGGATC 7487  
Qy 1680 GTCAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1739  
Db 7488 GTCAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7547  
Qy 1740 TGTCTCATGAGCGGATGCGAAGAGTGTATTTGCTGAGTATGATGATGATGATGATG 1799  
Db 7548 TGTCTCATGAGCGGATGCGAAGAGTGTATTTGCTGAGTATGATGATGATGATGATG 7607  
Qy 1800 GA 1801  
Db 7608 GA 7609

RESULT 2  
US-09-073-587-2  
Sequence 2, Application US/09073587  
Patent No. 6268552  
GENERAL INFORMATION:  
APPLICANT: Li, Yi  
TITLE OF INVENTION: Transgenic Seedless Plants  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
City: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/073,587  
FILING DATE: 06-MAY-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/045,725  
FILING DATE: 06-MAY-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Feiber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 4-97

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2211 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEICAL: NO  
AUTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 4..2205  
US-09-073-587-2

Query Match 94.2%; Score 1697.4; DB 3; Length 2211;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1726; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 64 ACAATGTCATTAAGGCGGATGATTTGACCGCGGTTTCCGATGCTTCTTGAACGA 123  
DB 1 ACAATGTCATTAAGGCGGATGATTTGACCGCGGTTTCCGATGCTTCTTGAACGA 60  
QY 124 GAAGCTCTAGGGGAGAGGATTAACCAAGCTCCAGTGGAGGCGGTTAGCT 183  
DB 61 GAAGCTCTAGGGGAGAGGATTAACCAAGCTCCAGTGGAGGCGGTTAGCT 120  
QY 184 TGCAAAAGGCTGCGATGCTGCTCCGAGATCTCAGTGGTGAAGAGTAGAGT 243  
DB 121 TGCAAAAGGCTGCGATGCTGCTCCGAGATCTCAGTGGTGAAGAGTAGAGT 180  
QY 244 CTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGGAGATCTTGAATCGAAACT 303  
DB 181 CTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGGAGATCTTGAATCGAAACT 240  
QY 304 TGGGCGGGGCAACAGAGTGTCTGTGGCAATGCACTTGGACACATTTGGATGAT 363  
DB 241 TGGGCGGGGCAACAGAGTGTCTGTGGCAATGCACTTGGACACATTTGGATGAT 300  
QY 364 TTCTCCGACACAACTATTCAGACCCCTGTTTTGCTGAGCGGTAAAGATGTCACCG 423  
DB 301 TTCTCCGACACAACTATTCAGACCCCTGTTTTGCTGAGCGGTAAAGATGTCACCG 360  
QY 424 ATTGATCTTAGTCAATTTGCTGCGCATTTCAATCTTAAGACTGCGGCTTTCGAACTTG 483  
DB 361 ATTGATCTTAGTCAATTTGCTGCGCATTTCAATCTTAAGACTGCGGCTTTCGAACTTG 420  
QY 484 CCAATGCGGCTGACGAGATGAGATGCAATGCAATGCGCTTACCGGTTTACATACCTT 543  
DB 421 CCAATGCGGCTGACGAGATGAGATGCAATGCAATGCGGCTTACCGGTTTACATACCTT 480  
QY 544 GAAGGGCGGCTGCAATTTGACATGTTAGTGTGGAACCTGATGCTGAAGGTTTCG 603  
DB 481 GAAGGGCGGCTGCAATTTGACATGTTAGTGTGGAACCTGATGCTGAAGGTTTCG 540  
QY 604 GCAAGTTCTTTCCACATGCACTTGTCTGACATGCAAGCCGTTTTTGAACCAATGT 663  
DB 541 GCAAGTTCTTTCCACATGCACTTGTCTGACATGCAAGCCGTTTTTGAACCAATGT 600  
QY 664 TCCGATGTTGAGGAGATGCGGCTTCTTCCGAGGATTTCTTAAGCCGGAAGTGGCGTTC 723  
DB 601 TCCGATGTTGAGGAGATGCGGCTTCTTCCGAGGATTTCTTAAGCCGGAAGTGGCGTTC 660  
QY 724 ATTGCGCTGCAATTTCCGACTGCTGTGGCAACGAAGTGTCTTCAATGCTGGGGTAGAC 783  
DB 661 ATTGCGCTGCAATTTCCGACTGCTGTGGCAACGAAGTGTCTTCAATGCTGGGGTAGAC 720  
QY 784 GATGTTTCAATATATGAGCAAGTATGCTGTGGAGGCAAGCTTTGGTCAATGCTTTC 843  
DB 721 GATGTTTCAATATATGAGCAAGTATGCTGTGGAGGCAAGCTTTGGTCAATGCTTTC 780

QY 844 AGGAGCGCTCTAGTGTGCGGCCGAATGAGGGCGATGCGATTTCTCTGCTGATTC 903  
DB 781 AGGAGCGCTCTAGTGTGCGGCCGAATGAGGGCGATGCGATTTCTCTGCTGATTC 840  
QY 904 TGCCTGTTTTTCTTCTCGAGCGTTACCGGCTGTCTTGTGATGAGCGGTTCCCAATCCC 963  
DB 841 TGCCTGTTTTTCTTCTCGAGCGTTACCGGCTGTCTTGTGATGAGCGGTTCCCAATCCC 900  
QY 964 GGCACATGACACTTACTTGTGCTTACCAAGGCTCCAAATCATGTGGAAGCCGGGACG 1023  
DB 901 GGCACATGACACTTACTTGTGCTTACCAAGGCTCCAAATCATGTGGAAGCCGGGACG 960  
QY 1024 CTGCCACCGAAGCTGTCCATCGCTTTTACCAAGGTTGGCGCTGTTCTTGAAGACGCT 1083  
DB 961 CTGCCACCGAAGCTGTCCATCGCTTTTACCAAGGTTGGCGCTGTTCTTGAAGACGCT 1020  
QY 1084 TTTTCATGAGGAGATATGTTGTTGGCTTGGCTGTGCTTATTAATCAAGGCTTGAATCA 1143  
DB 1021 TTTTCATGAGGAGATATGTTGTTGGCTTGGCTGTGCTTATTAATCAAGGCTTGAATCA 1080  
QY 1144 GGCACATTAAGTGGGCTCATGCTCTGGCAAAATTTGGCTGAACCGTTTCGGAGGAG 1203  
DB 1081 GGCACATTAAGTGGGCTCATGCTCTGGCAAAATTTGGCTGAACCGTTTCGGAGGAG 1140  
QY 1204 TCTCTCTTCAAGGATGAGAGGATCTTCTGGGACACATCTCTGCTGGTGGT-AAACA 1262  
DB 1141 TCTCTCTTCAAGGATGAGAGGATCTTCTGGGACACATCTCTGCTGGTGGTAAACA 1200  
QY 1263 TGGAGTTTCTTCAATGATTTGGGACCTTATCAAGTATGGAATAGGATCTGGCGGTT 1322  
DB 1201 TGGAGTTTCTTCAATGATTTGGGACCTTATCAAGTATGGAATAGGATCTGGC-GGGTT 1259  
QY 1323 TGGTCAAGTTTGAAGCGGTTTATTTAGATCTCTCGCTTGTGATCAAGCGATGTA 1382  
DB 1260 TGGTCAAGTTTGAAGCGGTTTATTTAGATCTCTCGCTTGTGATCAAGCGATGTA 1319  
QY 1383 AGAAATTCAGCGGATGTCCTTGAAGAAATCTGAGAACTTCCAGTCCGATGCGATCTGA 1442  
DB 1320 AGAAATTCAGCGGATGTCCTTGAAGAAATCTGAGAACTTCCAGTCCGATGCGATCTGA 1379  
QY 1443 AGTGTGTTAAGGTTGTCTGTGAGCCAGCGCATATGCTCAATGTTCAAGTCAAGCGATTCA 1502  
DB 1380 AGTGTGTTAAGGTTGTCTGTGAGCCAGCGCATATGCTCAATGTTCAAGTCAAGCGATTCA 1439  
QY 1503 GAAGGAAAACAAAATTAAGATTAAGGCTTAAAGACCGGATATCTGAACCTTATGATTA 1562  
DB 1440 GAAGGAAAACAAAATTAAGATTAAGGCTTAAAGACCGGATATCTGAACCTTATGATTA 1499  
QY 1563 GGTGTGTGTCACATCTGGAATGCAATATCCAACTCAGGCATTTGCTGACATGCGATAC 1622  
DB 1500 GGTGTGTGTCACATCTGGAATGCAATATCCAACTCAGGCATTTGCTGACATGCGATAC 1559  
QY 1623 CAATATTTTCAAGGACACAGTGAACCAAGCGGTTGTTAAACAGCATATGACAGATGCTC 1682  
DB 1560 CAATATTTTCAAGGACACAGTGAACCAAGCGGTTGTTAAACAGCATATGACAGATGCTC 1619  
QY 1683 AAAAATCTTCTGATGATGAAAGAAAATTTCTGTTTGAACCATATCTCTCCGTTGTGT 1742  
DB 1620 AAAAATCTTCTGATGATGAAAGAAAATTTCTGTTTGAACCATATCTCTCCGTTGTGT 1679  
QY 1743 CCTCATGAGCGGATGCAAAAGCAAGTATTTGCTGAGCTATGATTCGAGGATCGCA 1801  
DB 1680 CCTCATGAGCGGATGCAAAAGCAAGTATTTGCTGAGCTATGATTCGAGGATCGCA 1738

RESULT 3  
US-08-008-216-8  
; Sequence 8, Application US/08008216  
; Patent No. 536887  
; GENERAL INFORMATION:  
; APPLICANT: Slightom, Jerry L.  
; APPLICANT: Tepfer, David A.  
; TITLE OF INVENTION: Ri T-DNA Promoters

NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GRAY, CARY, AMES & FRYE  
STREET: 401 B Street, Suite 1700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-4297  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/008,216  
FILING DATE: 25-JAN-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/725,368  
FILING DATE: 22-APR-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhorst, Marlene W.  
REGISTRATION NUMBER: 36,740  
REFERENCE/DOCKET NUMBER: P1020US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 699-2700  
TELEFAX: (619) 236-1048  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2279 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Agrobacterium rhizogenes  
STRAIN: Strain A4  
IMMEDIATE SOURCE:  
LIBRARY: Convolulus arvensis plant cells  
CLONE: Clone 7  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..2279  
OTHER INFORMATION: /label= ORF8  
OTHER INFORMATION: /note= "Sequence ORF8 corresponds to bases 6609  
; OTHER INFORMATION: through 8888 of Seq. ID No. 536887 19."  
; US-08-008-216-8  
Query Match 6.8%; Score 121.6; DB 1; Length 2279;  
Best Local Similarity 45.7%; Pred. No. 1,56-32;  
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;  
QY 623 TCAGCTTGCTTACGACTGACGACCGTTTTCACCAATGTTCCGATAGTGAACGATCG 682  
DB 665 TTGACCTGCGGTACGACTATCAGGTTTCTCCAGAGTGCAGATGCCATGATGATGTGG 724  
QY 683 GCTTCTTCCGAGGAGATGTTCTTAAGCCGAAAGTGGGGTCATTGGCGCTGCATTTCCG 742  
DB 725 GGCATTATCCGGAAGGGGACACACTACCTAGTGTTCATCTGTAGAGGGGGGCTGTCTG 784  
QY 743 GACTCGGTGTGCAACGAACTGCTCATGCTGGGGTAGAGATGTACAAATATATGAG 802  
DB 785 GCCTTGTGCTGCCACGAACTACTTGGCGCTGCGTCGCAAGGAATACCTCTTTTCGATA 844  
QY 803 CAAGTATCGTGT---TGAGGCAAGCTTTGTCACATGCTTTTCAGGAGCGCTCTTAGTG 859  
DB 845 CCGTGTATGAGATCCGATGTTTGGGGCATCGCCGATGCCAAGCGGAGCGCTCACAGG 904  
QY 860 TCGTGGCCGAAATGGGGCGCATGCGATTTCTCTGCTGCATTTGCTTTTCTTCC 919  
DB 905 CTTTGAAGTGTGTGCTGTCTTCTTCCGCAACCACTTTGCTCTCATCTATC 964  
QY 920 TCGAGCGTTACGGGCTGTCTTCATGAGGCGGTTCCCAAAATCCCGGACAGTGACACTT 979

DB 965 TGATATAGTTAGTAATTCGCTCCAGCCTTCTTTCTTGTGCGGCAACGACACACAG 1024  
QY 980 ACTTGTCTTACCAAGCGCTCAATACATGTGAAAGCCGGGACGTGCCACCGAAGCTGT 1039  
DB 1025 CACTATATTTCCCGCAGAAACGCTACGATGCGACGGGGCAAGCTCCCGCGGGANTAT 1084  
QY 1040 TCGATCGCGTTTCAACGGTGGCGGCGCTTCTTGAAGACCGTTTTCATGACCGATA 1099  
DB 1085 TTACGGGGGACATGTCGGATGGAAGACACTACTTACCAAGGCTGTGAACGAAATGCA 1144  
QY 1100 TTGTGTGGCTTGGCTGCTGCTATTATCTACAGCGCTTGAATACAGACACTTAGTGAG 1159  
DB 1145 GGAAGCTGATGGCTCCGATGATATCTCTTCAATGTTGAAGAGCGCTGCGATGAGAG 1204  
QY 1160 CTGATATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGGCTCTTCTTACAGGA 1219  
DB 1205 CCTCAGAGACCGGACGCTTGGCTCCGAGAGTTCGGAATAATTCATTTTCATGCGGTTT 1264  
QY 1220 TAGAGAGATCTTCTGCGCACATCCCTCGTGCT---AAACATGAGTTTCTTC 1275  
DB 1265 TGGTCAGATCTTCAAGCTGTGTAATTCAGTCTGCTGCAAGGCAATGCAAAACACCCC 1324  
QY 1276 ATGATTGGGACCTATTCAAGCTAATGGAATAGATCTGCGGGGTTTGGTCCAGTTT 1335  
DB 1325 ATGATTTCAAGGCTTTGGGATATCTGAGTTGAGTATACGGCCGAGTTTGTCTC-TATTAC 1383  
QY 1336 GAAAGCGGTTTATTAGATCTCCGCTTGCTGATCAACGATATGAAAGAAATCAGCGG 1395  
DB 1384 AACGTGTGTTTCAACGATCCGAGACTGATTAATCAATGCTACAGAGGAGACAGCAT 1443  
QY 1396 ATGCGCCCTGAAGGATATCCAGAACTTCACGTCGATGCGATCTGAAGGTTAACGT 1455  
DB 1444 CTTCATTGTTGGGGTTCACTTTTGCAGGCTCTGATGCGCAATTAAATATTTCCAGAA 1503  
QY 1456 GTGCTGTGAGCCAGCGCATATGSCATGTTCAAGTCAGGCGCATTCGAAAGAAAGACA 1515  
DB 1504 AGCCATGCGAAGACAGACTCTGTTTGAATCCGTCGCTGGAATACCAAGAGGGCGGG 1563  
QY 1516 AAAATPAAGATTAAGCTTAAAGCGGGAATATCTGAACCTTATGATAGAGTGTGTACA 1575  
DB 1564 AGATTGAAGTATGCTTGAACACGCTCATTCGCGTGTGTTTGAACAGGCTCATTTGCG 1623  
QY 1576 TCTGAGCTGGCAATATCAACTCAGGCACTGCGCTGACATGCGATACCATATTTTCAG 1635  
DB 1624 GCGAGTCTGAGGCGCCTACAGTTGATTAACAAGCTGCGGGGATAGACTTCTTCAGC 1683  
QY 1636 GCACAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATCTGCAAAACTCTTCTG 1695  
DB 1684 TACAATATGAAACCGCGCGTGGAAACTGTCGCGCTGTCAATTCAGACCTCTTCATG 1743  
QY 1696 ATGACTGAACGAAATTTCTGTTAGACATATCTCCGCTTTGTGCTCATGACGGG 1755  
DB 1744 GTCCAGCAAGCAAAAGTTTGGGTTAACTCCGCGATCCAGCAAGTATATGACCGATGG 1803  
QY 1756 ATCGCAAAAGCAAGTATTTGCTGACTAGCTATGAGTGCAGAGAT 1797  
DB 1804 CTGTGCTGAGCTGTGTTGACATGCAATGCGCAGCT 1845  
RESULT 4  
US-08-459-569-8  
; Sequence 8, Application US/08459569  
; Patent No. 5543501  
; GENERAL INFORMATION:  
; APPLICANT: Slightom, Jerry L.  
; APPLICANT: Tepfer, David A.  
; TITLE OF INVENTION: R1 T-DNA Promoters  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GRAY, CARY, AMES & FRYE  
; STREET: 401 B Street, Suite 1700  
; CITY: San Diego

```

1 STATE: California
2 COUNTRY: USA
3 ZIP: 92101-4297
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: Patentin Release #1.0, Version #1.25
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/459,569
13 FILING DATE: 02-JUN-1995
14 CLASSIFICATION: 530
15
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 08/008,216
18 FILING DATE: 25-JAN-1993
19 APPLICATION NUMBER: US 06/725,368
20 FILING DATE: 22-APR-1985
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Barnhorst, Marlene W.
23 REGISTRATION NUMBER: 36,740
24 REFERENCE/DOCKET NUMBER: P1020US1
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (619) 699-2700
27 TELEFAX: (619) 236-1048
28
29 INFORMATION FOR SEQ ID NO: 8:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 2279 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: double
34 TOPOLOGY: linear
35
36 MOLECULE TYPE: DNA (genomic)
37
38 ORIGINAL SOURCE:
39 ORGANISM: Agrobacterium rhizogenes
40 STRAIN: Strain A4
41 IMMEDIATE SOURCE:
42 LIBRARY: Convolvulus arvensis plant cells
43 CLONE: Clone 7
44
45 FEATURE:
46 NAME/KEY: misc_feature
47 LOCATION: 1..2279
48 OTHER INFORMATION: /label= ORF8
49 OTHER INFORMATION: /note= "Sequence ORF8 corresponds to bases 6609
50 Patent No. 5543501
51 OTHER INFORMATION: through 8888 of Seq. ID No. 5543501 19. "
52 US-08-459-569-8
53
54 Query Match 6.8%; Score 121.6; DB 1; Length 2279;
55 Best Local Similarity 45.7%; Pred. No. 1,5e-32;
56 Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;
57
58 QY 623 TCGACTTGCTCTACGACTGCAGACCGTGTCTTTTGAACCATGTGTTCCATATAGTAGGACGATCG 682
59 |||||
60 Db 665 TTGACCTGCGGTAGCAGTACATCAGTTTCCCTCCAGAGATGGCATGCCATATAGTGTG 724
61
62 QY 683 GCTTCTTTCCGAGAGATTTCTTAAGCCGAAAGTGGCGGTCAATTGGCGCTGGCATTTCCG 742
63 |||||
64 Db 725 GGCATTTATCCGGAAGGCCACACACTACCTAGTGTGCCATCGTAGAGAGCGGCGTGTCTG 784
65
66 QY 743 GACTCGTGTCGCAACGCAACTGCTTACGTGCGGGGTAGACGATTTACATATATATAGAG 802
67 |||||
68 Db 785 GCCTGTGTCTGCCACACAACTACTTGGCGGTGGGCTGAAGGAATCACTCTTTTCGATA 844
69
70 QY 803 CAACTGATCTGT---TGAGGCAAGCTTTGGTGCATGCTTTTCAGGAGCGCTCTAGTG 859
71 |||||
72 Db 845 CCGTTGATGAATCCGTAGTTTGGGGCATGCCGATGCCCAACGGCGACCTCACAG 904
73
74 QY 860 TCGTGCGCAAAATGGGGCGATGCGATTCTCTCTGTGCATCTGCTTGTGTTTTCTTCC 919
75 |||||
76 Db 905 CTTGACGCTCGTGGTGTCAATGCTTTCTCCGCCAACCAACTTGGCTGTCACTATATC 964
77
78 QY 920 TCGAGCGTTAGGCGCTGCTTCGATGAGAGCGCGTTCCCAATCCGGGCAAGTGCAGACTT 979
79 |||||
80 Db 965 TGGATTAAGTTAGAAATCCGTCCAGCTTTCGTTTCTTGTGCGGGCAAGCACACAG 1024

```

Qy	980	ACTTGGCTCAACGAAGGGCTCAATAACATGTGGAAGCGGGGACAGCTGCACCCGAACCTGT	1039
Db	1025	CACATAATTTCCGCGCAAGAACGCTACGCAATGGCACGCGGGGGACAGCTCCGCGGGGATAT	1084
Qy	1040	TCCATCGCGTTTACMACGGTTGGCGCTGCTCTTGAAGACGCGTTTTCATGAGCGAGATA	1099
Db	1085	TTGACGGGGTACATGTCGGATGGAAAGACACTACTCTACAAAGGGGTGGAAACGGAAATGGCA	1144
Qy	1100	TTGTGTTGGCTTCGCTCTGTGGCTATTACTAGGCTTTGAANAATCAGACACATTAAAGTCGG	1159
Db	1145	GGAGACATGATGGCTCCCATGGATAATCTCTTTCATGTGTGAAGACCGTCGTGATGTAAG	1204
Qy	1160	CTCATGACCTCGGCAAAATTTGGCTGAACCGTTTGGGAGAGGAGTCCCTTCTTCACAGGA	1219
Db	1205	CCTGAGAAAGCAGGACAGCTTTGGCTCCGAGAGTTTCGAAAAATTCATCTTTCATGCCCTTT	1284
Qy	1220	TAGAGAGGATCTTTCCTGGGACACATCTCTCGTGGT---AAACATGAGATTTCCTC	1275
Db	1265	TGTCGAGATCTTCAGCTGTGTGTAATTGCAATCCTGGTGGCAAGGCATGGCAAAACATCCC	1324
Qy	1276	ATGATTTGGACCTATTACAGCTAATGGGAATAGGATTTGGCGGGGTTTGGTCCAGTTTTT	1335
Db	1325	ATGATTTTCGAGGCTTTCGGGATACTGAGTTGGGATACGGCCAGATTTTCCTTC-TAATTAC	1383
Qy	1336	GAAAGCGGGTTTATATGATCTCTCCGCTGGTCAATCAACGGAATATGAAGAAATTCAGCGG	1395
Db	1384	AACGTGTTGTTTTCACAGATCTCTGGACTGATTAATCAATGGCTACAGAGAGACACAGAT	1443
Qy	1396	ATGTGCCCTGAGAGGAATCTCAGAACTTTCACAGTCGATTCGCATCTGAAGTGGTTAACGT	1455
Db	1444	CTTCTCATTTGTTGGGGTTCAACTTTTTCAGAGCTCTGATGCGCATTTGAATATTCAGAAA	1503
Qy	1456	GTCCTGTGAGCCAGGCGCATATGCAATGTTCAAGTCAAGGCGCATTCAGAAAGAAAGACA	1515
Db	1504	AGCCATCGGAAGCACACACTCTGTTTGGATCCGCTGGTGAATAGCCAAAGAGGGCGGG	1563
Qy	1516	AAATAAAGATTAAGGCTTAAAGAGGGGATATCTGAACCTTATGATTAAGTGGTGCACA	1575
Db	1564	AGATTTGAAGGTATGCTTGAAACACGCGTCAATTCGGGTGTTTGAACCGATCATCATTTGGC	1623
Qy	1576	TCTGGACTCGCAAAATATCAACTCAGGACTTGCCTGACATGCGATACCAATATTTTTCAG	1635
Db	1624	GGCAGTCTGAGGCGCGCTACAGTTGATTAACAGACTGGCCGGGGATGGAATTCCTTCAGC	1683
Qy	1636	GCACCACTGAACCAAGCGGTTGATTAACGCCATATGACAGAGATCGTAAAACTTTCTCTG	1695
Db	1684	TACATATTCGAACCCGCGCTGGAAATCTGCTGCGGCTGTCAATTCAGCACTTTCATG	1743
Qy	1696	ATGACTGAAGAAATCTGTTAGACATATCTCCCGCTTGTTGTCTCATGAGCGGG	1755
Db	1744	GTCAGAAAGCAAAAGTTTGGGTTAACTCCGGCATCCACAGACTGATATGAGACCGATGGG	1803
Qy	1756	ATGCAAAAGCAGTGTATTCCTGAGACTATGAGTCCGAGAT	1797
Db	1804	CTTGTCGCTGAGCTGTGTTGCATGTACATGGAATCCGCACT	1845
RESULT 5			
US-08-458-831-8			
: Sequence 8, Application US/08458831			
: Patent No. 5824866			
: GENERAL INFORMATION:			
: APPLICANT: Slightom, Jerry L.			
: APPLICANT: Tepier, David A.			
: TITLE OF INVENTION: RI T-DNA Promoters			
: NUMBER OF SEQUENCES: 19			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: GRAY, CARV, AMES & FRYE			
: STREET: 401 B Street, Suite 1700			
: CITY: San Diego			
: STATE: California			
: COUNTRY: USA			

ZIP: 92101-4297  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,831  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,216  
FILING DATE: 25-JAN-1993  
APPLICATION NUMBER: US 06/725,368  
FILING DATE: 22-APR-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhorst, Marlene W.  
REGISTRATION NUMBER: 36,740  
REFERENCE/DOCKET NUMBER: P1020US1  
TELEPHONE: (619) 699-2700  
TELEFAX: (619) 236-1048  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2279 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Agrobacterium rhizogenes  
STRAIN: Strain A4  
IMMEDIATE SOURCE:  
LIBRARY: Convolvulus arvensis plant cells  
CLONE: Clone 7  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..2279  
OTHER INFORMATION: /label= ORF8  
OTHER INFORMATION: /note="Sequence ORF8 corresponds to bases 6609  
Patent No. 5824866  
OTHER INFORMATION: through 8888 of Seq. ID No. 5824866 19. "  
US-08-458-831-8  
Query Match 6.8%; Score 121.6; DB 1; Length 2279;  
Best Local Similarity 45.7%; Pred. No. 1.5e-32;  
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;  
QY 623 TCAGCTGCTCTACGACTGAGACCGTTTTCACCAATGTTCCGATAGTGAAGCATCG 682  
DB 665 TTGACCTGGGTAAGTATACAGTTTCTCCAGAGTGGAGTCCCATGATGATGCTGG 724  
QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGTCAATTGGCGTGGCATTTCCG 742  
DB 725 GGCATTATCCGGAAGGCGCACACTACCTAGTGTGTCATGTGAGGAGGCGGCTGTCTG 784  
QY 743 GACTCGGTGGGCAAGCACTGCTTCATGCTGGGGTAGAGCATTAATATATAGAG 802  
DB 785 GCTTGTGTGTCACAGAACTACTTGCGCTGCGTCAAGAAATACCTTTTCGATA 844  
QY 803 CAAGTATCGTGT--TGAGAGCAAGCTTTGTGCATGCTTTCAGGAGCGTCTTAGTG 859  
DB 845 CCGTGTATGAGATCCGTAAGTTTGGGCGATCGCCGATGCCAAAGGCGACGCTACAGG 904  
QY 860 TCGTGCCGAAATGGGCGCATGCGATTTCTCTGCTGCATTTGCTTTTCTTCC 919  
DB 905 CCTGAGCGTGTGCTGATGCTTTCCTCCGCAACCACTTGGCTGTCACTATATC 964  
QY 920 TCGAGCGTTACGCGCTCTTTCATGAGGCGCTTCCCAATCCGCGACAGTGCACCTT 979  
DB 965 TGGATTAAGTTTGAATTCGCTCCAGCTTCTTCTTGGCCGGAACGACACACAG 1024  
QY 980 ACTTGCTACCAAGGCGCTCAATATACATGTAAGCAAGCGGCGACGCTGCCACCGAAGCTGT 1039

DB 1025 CACTATATTTCCGCCAGAAACGCTACGATGCGACCGGGGCAAGCTCCGCCGGGATAT 1084  
QY 1040 TCCATGCGCTTTACACGAGTTGGCGGCTTCTTGAAGACGGTTTTCATGACCGATA 1099  
DB 1085 TTGAGGGGTAAGTGTGGAAGAGACACTACTTCAACAAAGGTTGAAACGGAATGCA 1144  
QY 1100 TTGTGTGGCTTGGCGCTGTGCTATTACTAGCGCTTGAATCAGGACATTAAGTGGG 1159  
DB 1145 GAGACTGATGCTCCGATGATATCTTTTCAATGTTAAAGAGCGCTCGTGTGATGAAG 1204  
QY 1160 CTCATATCTCTGGCAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTTTCAGGGA 1219  
DB 1205 CCTCAAGACGCGCAGCTTGTGGCTCCGAGAGTTCCGAAATTCATTTTCAAGCGCTT 1264  
QY 1220 TTAGAGGATCTTTCGGGACACATCTCTGGTGT--AAATGAGTATTCTC 1275  
DB 1265 TGGTCAAGATCTTCACTGTGTAAATTCAGTCTGTGGCAAGCATGGCAAAACCCC 1324  
QY 1276 ATGATTGGACCTATTCAAGCTAATGGAATAGATCTGGCGGGGTTTGGTCCAGTTT 1335  
DB 1325 ATGATTTCAAGGCTTTCGGGATCTAGGTTGGGATACGCGCAGTTTGTCTC-TATTAC 1383  
QY 1336 GAAAGCGGTTTATAGATCTCCGCTTGTGTCATCAAGATATGAAGAAATCAGCGG 1395  
DB 1384 AAGGTGTGTTTCAACGATCTCGACGTGATATCAATGCTACGAGGAGACAGCAT 1443  
QY 1396 ATGTGCCCTGAAGGATCTCAGAACTTCACGCTGGATGCGCATCGAAGGTGTTACGGT 1455  
DB 1444 CTTCATTATGTTGGGTTCACTTTTGACAGCTCTATGCGCATTAATATATTCAGAAA 1503  
QY 1456 GTGTCTGTGAGCGCAGCATATGCAATGTTCAAGTCAGGCGCATTAAGAGAAAGACA 1515  
DB 1504 AGCCATGCGAAAGACAGACTCTGTTTGAATCCGTCGTGTGATTAACCAAGAGGGCGG 1563  
QY 1516 AAAATAAGATAAGCTTAAAGCGGATATCTGAATTTATGATTAAGTGTGTCACA 1575  
DB 1564 AGATTGAAGTATGCTTGAACACAGGCTCATTCGCTGTTTGAACAGGCTCATATGCGC 1623  
QY 1576 TCTGAGCTGCAAAATATCAACTCAGGCAATGCTGTCATGATGCATATATTTTCAG 1635  
DB 1624 GGCAGTGTGAGGCGCGCTACAGTTGATTAACAGACTGCGGGATGATGATCTCTTACG 1683  
QY 1636 GCACAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATCTCAAAACTCTTCTG 1695  
DB 1684 TAAATATGAAACCGCGCGTGGAACTGCTGCGGCTGTCAATTCAGACTTTTATG 1743  
QY 1696 ATGACTGAACGAAATTTGTGTTAGACCATATCTCCGCTTGTGCTCATGACGGG 1755  
DB 1744 GTCACGAAGCAAAATTTGGGTTAACTCCGCGCATCCAGCACTGATGATGACCGATGGG 1803  
QY 1756 ATGCAAAAGCACTGATATGCTCGGACTATGATCCAGGAT 1797  
DB 1804 CTGTCCGTGAGCTGTGTGATTAACATGCAATCCGACGT 1845  
RESULT 6  
US-08-008-216-19  
Sequence 19, Application US/08008216  
Patent No. 5366887  
GENERAL INFORMATION:  
APPLICANT: Slightom, Jerry L.  
APPLICANT: Tephner, David A.  
TITLE OF INVENTION: R1 T-DNA Promoters  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GRAY, CARY, AMES & FRYE  
STREET: 401 B Street, Suite 1700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-4297  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/008,216  
FILING DATE: 25-JAN-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/725,368  
FILING DATE: 22-APR-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhorst, Marlie W.  
REGISTRATION NUMBER: 36,740  
REFERENCE/DOCKET NUMBER: P1020US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 699-2700  
TELEFAX: (619) 236-1048  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21126 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Agrobacterium rhizogenes  
STRAIN: STRAIN A4  
IMMEDIATE SOURCE:  
LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS  
CLONE: CLONE 7  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (937..2262)  
OTHER INFORMATION: /label= ORF1SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (2649..3458)  
OTHER INFORMATION: /label= ORF2SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (4041..4400)  
OTHER INFORMATION: /label= ORF4SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (4041..4400)  
OTHER INFORMATION: /label= ORF4SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (4607..4918)  
OTHER INFORMATION: /label= ORF5SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (5143..6216)  
OTHER INFORMATION: /label= ORF6SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (5071..5643)  
OTHER INFORMATION: /label= ORF7SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (6609..8888)  
OTHER INFORMATION: /label= ORF8SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (6576..6830)  
OTHER INFORMATION: /label= ORF9SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (9748..10044)  
OTHER INFORMATION: /label= ORF10SUBSEQUENCE  
FEATURE:

NAME/KEY: misc feature  
LOCATION: complement (10509..11282)  
OTHER INFORMATION: /label= ORF11SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 12466..13002  
OTHER INFORMATION: /label= ORF12SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 13723..14319  
OTHER INFORMATION: /label= ORF13SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 15659..16210  
OTHER INFORMATION: /label= ORF14SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (16517..17545)  
OTHER INFORMATION: /label= ORF15SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (17737..18189)  
OTHER INFORMATION: /label= ORF16SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (18177..18743)  
OTHER INFORMATION: /label= ORF17SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: complement (19031..19390)  
OTHER INFORMATION: /label= ORF18SUBSEQUENCE  
US-08-008-216-19

Query Match 6.8%; Score 121.6; DB 1; Length 21126;  
Best Local Similarity 45.7%; Pred. No. 8,4e-32;  
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;

QY 623 TCGACTTGTCTACGACGACGACCGTTTTCGACCAATGTTCCGATGAGCGATCG 682  
DB 7273 TTGACCTGGGTAGCGACTACAGTCTTCTCCAGAGTGCATGCGCCATGATGATGTGG 7332  
QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCGGTCATTGGCGCTGCAATTCG 742  
DB 7333 GGCATTATCCGAAGCGCACCACTACCTAGTGTTCGATCGTAGAGAGGGGCTGTCTG 7392  
QY 743 GACTCGTGTGGCAAGCAATGCTTCATGCTGGGCTAGACGATGTTACATATATGAAG 802  
DB 7393 GCCTTGTTGCTGCCACAGAACTACTTGGCGCTGCAAGAAATCACTTTTCGATA 7452  
QY 803 CAAGTATGCTGT--TGAAGGCAAGCTTTGTCATGCTTTTCAGGAGCGTCTTAGTG 859  
DB 7453 CCGTTATATGATGATCCGTAGTCTTTGGGCAATCGCCGATGCCAAGCGCGCTCACAGG 7512  
QY 860 TCGTGCCCGAAATGGGGGCGATGCGATTTCTCTGCTGATTCGCTGTTTCTTC 919  
DB 7513 CTTGACGTCGTTGCTGTCATGCTTTCTCCGCCAACAACCTTCTGCTGCTATATC 7572  
QY 920 TCGAGCGTTACGCGCTGTCTTGATGAGCGCGTTCCCAATCCCGCACAGTGCACATT 979  
DB 7573 TGGATTAAGTTTAAGATTCGTCAGCTTCGTTTCTTGTGCGGCAAGCACACAG 7632  
QY 980 ACTTGTCTACCAAGCGCTCAATATGATGGAAGCGGGGAGCTGCCACCAAGCTGT 1039  
DB 7633 CACTATATTTCCGCCAAGACGCTAGCAATGCGACCGGGGCAAGCTCCGCGGGGATAT 7692  
QY 1040 TCCATGCGCTTAAACAAGGTTGGCGGCTTCTTGAAGAGCGGTTTTCATGACGAGATA 1099  
DB 7693 TTACAGGCGTACATGTCGATGGAAGCACTACTCTACCAAGGCTGTAAGCAATGCGCA 7752  
QY 1100 TTGTGTGCTTTCGCTGCTGCTATTAATCAAGCTTGAATTCAGCAATTAAGTGGG 1159  
DB 7753 GGAGACTGATGCGCTCGATGATATCTCTTCAATGTTGAAGAGCGTCTGCTGATGAAG 7812

```

QY      1160  CTCATGACTCTGGCAAAATTTGGCTGGAACCGTTTCGGGAGGAGGAGTCCTTCTTCAGGGA 12119
Db      7813  CCTGAGAAGCAGCGGACGTTGGCTCCGAGAGTTGGAAAATTCATCTTCATCGCGTT 78722
QY      1220  TAGAGAGATCTTCTGGGCAACATCTCTCGTGGT- - -AAACATGAGTTTCTC 1275
Db      7873  TGTGTCGAGATCTTCAGCTGTGTGTAATTGATGCTCTGGTGGCAAGCATGCAACACCCC 79322
QY      1276  ATGATTTGGACCTTATTCAGCTAATGGAAATAGCATCTGGGGGGGTTTGGTTCAGTTT 13355
Db      7933  ATGATTTTCGAGGCTTTTCGGGATCTGAGTGTTGGGATACGGCCGAGTTTCGTC- TATTAC 79911
QY      1336  GAAAGCGGGTTTATTGAGATCTCCGCTGGTGCATCAACGGATATGAAGAAATACGGG 13955
Db      7992  AACGTGTTGTTTTCACGATCTTGAGCTGGAATTATCAATGCTACGAGAGACACGAT 80511
QY      1396  ATGTGCCCTGAAGAAATCTCAGAACTTCACGTCGGATCGCATCTGAAGTGTTAACGCT 14555
Db      8052  CTTTCTATATGTTGGGGTTCAACTTTTTCAGGCTCTGATGGCGCATTTGAATTCAGAAA 81111
QY      1456  GTGTCCTGAGCCAGCCGATATGCCATGTTCAAGTCAGGGCGATTCAGAGAAAAGACA 15155
Db      8112  AGCCATGGAAAGACGACGACTGTTTGTGATCCCGTCTGGAATAAGCCAGAGAGGCGGG 81711
QY      1516  AAATTAAGATTAAGGCTTAAGAGGGGGATCTGAACCTTATGATTAAGGGTGTGCACA 15755
Db      8172  AGATTGAAGGTATCTTGAAACACGATCATTCGGGTGTTTTCACAGGTATCATATGGC 82311
QY      1576  TCTGACTCGCAAAATATCAACTCAGGATTCCTGACATGCGATACCAATATTTTTAG 16355
Db      8232  GCGAGTGTGAGGCGCGCTACAGTTGATPAACAGACTGGCGGGGATGAGACTTCTTACG 82911
QY      1636  GCACCACTGAACCAAGCGGTGATPAACAGCATATGACAGATGCTCAAACTTTCTG 16955
Db      8292  TACATATCGAACCCGCGCTCGAAACTCGTCCGCTGTCAAATTCAGCACTCTTCATG 83511
QY      1696  ATGATGACGAATAAATTCGTGTAGTACCATATCTCTCCGCTTGTGTCTCATGAGACGG 17555
Db      8352  GTCAAGCAAGCAAAAGTTTGGGTTAACTCCGCAATCCACGAGATGATATGAGACCGATGG 84111
QY      1756  ATCCCAAAAGCAGTGTATTTGCTGGACTATAGTCGACAGAT 1797
Db      8412  CTTGTCGTCGAGCTGTGTTGCAATTGACATCGAATCGCAGCT 8453

RESULT 7
US-08-459-569-19
; Sequence 19, Application US/08459569
; Patent No. 5543501
; GENERAL INFORMATION:
; APPLICANT: Slighton, Jerry L.
; APPLICANT: Tepfer, David A.
; TITLE OF INVENTION: R1 T-DNA Promoters
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY, CARY, AMES & FRYE
; STREET: 401 B Street, Suite 1700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-4297
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,569
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,216
; FILING DATE: 25-JAN-1993

```

APPLICATION NUMBER: US 06/725,368  
FILING DATE: 22-APR-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: Barthorst, Maetrie W.  
REGISTRATION NUMBER: 36,740  
REFERENCE/DOCKET NUMBER: P1020US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 699-2700  
TELEFAX: (619) 236-1048  
IMMEDIATE SOURCE:  
LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS  
CLONE: CLONE 7  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (937..2262)  
OTHER INFORMATION: /label= ORF1SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (2649..3458)  
OTHER INFORMATION: /label= ORF2SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3726..4799  
OTHER INFORMATION: /label= ORF3SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (4607..4918)  
OTHER INFORMATION: /label= ORF5SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 5143..6216  
OTHER INFORMATION: /label= ORF6SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (5071..5643)  
OTHER INFORMATION: /label= ORF7SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6609..8888  
OTHER INFORMATION: /label= ORF8SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 9748..10044  
OTHER INFORMATION: /label= ORF10SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (6576..6830)  
OTHER INFORMATION: /label= ORF9SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 12466..13002  
OTHER INFORMATION: /label= ORF12SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature

LOCATION: 13723..14319  
OTHER INFORMATION: /label= ORF13SUBSEQUENC  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 15659..16210  
OTHER INFORMATION: /label= ORF14SUBSEQUENC  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 16517..17545  
OTHER INFORMATION: /label= ORF15SUBSEQUENC  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 17737..18189  
OTHER INFORMATION: /label= ORF16SUBSEQUENC  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 18177..18743  
OTHER INFORMATION: /label= ORF17SUBSEQUENC  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 19031..19390  
OTHER INFORMATION: /label= ORF18SUBSEQUENC  
US-08-459-569-19

Query Match 6.8%; Score 121.6; DB 1; Length 21126;  
Best Local Similarity 45.7%; Pred. No. 8.4e-32;  
Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;

QY 623 TCGACTTGTCTTACGACTGACAGCCGTTTGTGACCAATGTTCCGATAGTGAACGATCG 662  
DB 7273 TTGACCTGCGGATGACACTATCAGGTTTCTCCAGAGTGGATGCCATGATGTTGTG 7332  
QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCGAAAGGGGGTCAATGSGGCTGGCATTTCCG 742  
DB 7333 GGCATTATCCGAAAGGCGACACTACTAGTGTGTCATGTAAGAGGCGGCTGTCTG 7392  
QY 743 GACTGTGTGTGCAAGAACTGCTTCATGCTGGGTAGACATGTTACATATATGAG 802  
DB 7393 GCGTGTGTGCGCCACAGAACTACTTGGCGTGGCGTCAAGAAATACCTTTTCGATA 7452  
QY 803 CAAGTATCGTGT--TGAAGGCAAGCTTTGTGTCAATGCTTTTCAGGAGCGTCTTAAG 859  
DB 7453 CCGTGTGTGAGATCCGATGTTTGGGGCATGCGCATGCCAAAGCGGAGCTCACAGG 7512  
QY 860 TCGTGGCCGAAATGCGGCGCATGCGATTTCTCTGTCGATCTGCTTTTCTTCTCC 919  
DB 7513 CTTGAGT 7572  
QY 920 TCGAGCGTACGCGCTGTCTTTCATGAGCGCGTCCCAATGCCGAGACAGTGCACATT 979  
DB 7573 TGGATTAAGTTAAGATTCCTGCTCCAGCTTGTCTTGTGCGGCAACACTTGTCTCATATC 7632  
QY 980 ACTTGTCTACCAAGGCGTCCATATCATGTGAAAAGCGGCGAGCTGCCAGAGCTGT 1039  
DB 7633 CACTATATTTCCGCAAGAAAGCTACGATGCGAGCGGCGCAAGCTCCCGGGGATAT 7692  
QY 1040 TCGATCGCGCTTCAACAGGTTGGCGTCCGTTCTTGAAGAGCGTTTTCATGAGCGAGATA 1099  
DB 7693 TTACGCGGTAACATGTGTGATGAAACACTACTCTTCAAGAGGTGTGAACGGAATGCA 7752  
QY 1100 TTGTGTGTGCTGCGCTGTGCTATTTACTCAGGCTTGAATCAGACACATTAGTGGG 1159  
DB 7753 GAGACCTGATGCTCCGATGATATCTTTTCATGTTGAAGAGCGTGTGCTGATGAAG 7812  
QY 1160 CTCATGACTCTGCAAAATTTGCTGAACCGTTTCGAGAGGAGTCTTCTTTCAGGGA 1219  
DB 7813 CTCAGAAAGACGCGACCTTTGGCTCGAAGTTCGAAATTCATCTTTCATGCGGTTT 7872  
QY 1220 TAGAGAGATCTTTTGTGGGACACATCTCTGTGTGT---AAACATGAGTTTCTTC 1275  
DB 7873 TGTGTGAGATCTTCAAGTGTGTGTAATTCAGTCTCTGTGTGCAAGGACATGCAACACCCC 7932  
QY 1276 ATGATTGGAGCTATTCAGCTAATGGAATAGGATCTGCGGGGCTTGTGTCCAGTTT 1335

DB 7933 ATGATTTCGAGGCTTTCGGGATCTAGGTTGGGATACGGCCGAGTTTCTCC-TATTAC 7991  
QY 1336 GAAAGCGGTTTATTAGATCTCCGCTTGCTGATCAAGGATATGAAGAAATCAGCGG 1395  
DB 7992 AAGGTGTGTGTTTCAACGATCTCGAGCTGATTTATCAATGGCTACAGAGGAGCGACAT 8051  
QY 1396 ATGTGCGCTGAAGGAATTCAGAACTTCCACGTCGATGCGATCTGAAGTGTTAACGCT 1455  
DB 8052 CTTTCTATTGTGTGGGTTCACTTTTTCAGGCTCTGATGCGCATTAATATTTCCAGAAA 8111  
QY 1456 GTGTCTGTAGCCAGCGCATATGCTTCAAGTCAAGGCGCATTCAGAAAGAAAGACA 1515  
DB 8112 AGCCATGCGAAAGACAGACTCTGTTTGTATCCGTCGTGGAATPACCAAGAGGCGCGG 8171  
QY 1516 AAAATTAAGTAAAGCTTAAAGCGGGAATATCTGAACCTTATATAGTGTGTGTACA 1575  
DB 8172 AGATTGAAGTAAAGCTTGAACACGCTCATTCGCGTGTGTTTGAACAGGTCATCATTTGGC 8231  
QY 1576 TCTGACTGCAAAATATCCAACTCAGGCAATGCTGACATGCGATCAATATTTTCAG 1635  
DB 8232 GCGAGTGTGAGGCGGCTTCACTGATTAACAGCTGCGGAGATGAGACTTCTTCAGC 8291  
QY 1636 GCACAGTGAACCAAGCGGTTGATTAACAGCATATGACAGATCTCAAAACTCTTCTG 1695  
DB 8292 TACAATATGAACCCGCGCTCGGAAACTGTCTGCGGCTGTCAATTCAAGACTCTTCATG 8351  
QY 1696 ATGACTGAACGAAATTTCTGTGTTAGACATATCTTCCGCTTGTGTCTCATGACGGG 1755  
DB 8352 GTCCACAGCAAAAGTTTGGGTTAACTCCGCGATCCAGCACTGATATGAGCCGATGG 8411  
QY 1756 ATGCAAAAGCACTGATTCCTGCTGACTATGAGTGCAGGAT 1797  
DB 8412 CTTGCTCGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8453

## RESULT 8

US-08-458-831-19  
; Sequence 19, Application US/08458831  
; Patent No. 582486  
; GENERAL INFORMATION:  
; APPLICANT: Slightom, Jerry L.  
; APPLICANT: Tepfer, David A.  
; TITLE OF INVENTION: R1 T-DNA Promoters  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; STREET: 401 B Street, Suite 1700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-4297  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/458,831  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 800  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/008,216  
; FILING DATE: 25-JAN-1993  
; APPLICATION NUMBER: US 06/725,368  
; FILING DATE: 22-APR-1985  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barnhorst, Marlene W.  
; REGISTRATION NUMBER: 36,740  
; REFERENCE/DOCKET NUMBER: P1020US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 699-2700  
; TELEFAX: (619) 236-1048

1 INFORMATION FOR SEQ ID NO: 19:  
2 SEQUENCE CHARACTERISTICS:  
3 LENGTH: 21126 base pairs  
4 TYPE: nucleic acid  
5 STRANDEDNESS: single  
6 TOPOLOGY: linear  
7 MOLECULE TYPE: DNA (genomic)  
8 HYPOTHETICAL: NO  
9 ORIGINAL SOURCE:  
10 ORGANISM: Agrobacterium rhizogenes  
11 STRAIN: STRAIN A4  
12 IMMEDIATE SOURCE:  
13 LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS  
14 CLONE: CLONE 7  
15  
16 FEATURE:  
17 NAME/KEY: misc feature  
18 LOCATION: complement (937..2262)  
19 OTHER INFORMATION: /label= ORF1SUBSEQUENCE  
20  
21 FEATURE:  
22 NAME/KEY: misc feature  
23 LOCATION: complement (2649..3458)  
24 OTHER INFORMATION: /label= ORF2SUBSEQUENCE  
25  
26 FEATURE:  
27 NAME/KEY: misc feature  
28 LOCATION: 3726..4799  
29 OTHER INFORMATION: /label= ORF3SUBSEQUENCE  
30  
31 FEATURE:  
32 NAME/KEY: misc feature  
33 LOCATION: complement (4041..4400)  
34 OTHER INFORMATION: /label= ORF4SUBSEQUENCE  
35  
36 FEATURE:  
37 NAME/KEY: misc feature  
38 LOCATION: complement (4607..4918)  
39 OTHER INFORMATION: /label= ORF5SUBSEQUENCE  
40  
41 FEATURE:  
42 NAME/KEY: misc feature  
43 LOCATION: 5143..6216  
44 OTHER INFORMATION: /label= ORF6SUBSEQUENCE  
45  
46 FEATURE:  
47 NAME/KEY: misc feature  
48 LOCATION: complement (5071..5643)  
49 OTHER INFORMATION: /label= ORF7SUBSEQUENCE  
50  
51 FEATURE:  
52 NAME/KEY: misc feature  
53 LOCATION: 6609..8888  
54 OTHER INFORMATION: /label= ORF8SUBSEQUENCE  
55  
56 FEATURE:  
57 NAME/KEY: misc feature  
58 LOCATION: complement (6576..6830)  
59 OTHER INFORMATION: /label= ORF9SUBSEQUENCE  
60  
61 FEATURE:  
62 NAME/KEY: misc feature  
63 LOCATION: 9748..10044  
64 OTHER INFORMATION: /label= ORF10SUBSEQUENCE  
65  
66 FEATURE:  
67 NAME/KEY: misc feature  
68 LOCATION: complement (10509..11282)  
69 OTHER INFORMATION: /label= ORF11SUBSEQUENCE  
70  
71 FEATURE:  
72 NAME/KEY: misc feature  
73 LOCATION: 12466..13002  
74 OTHER INFORMATION: /label= ORF12SUBSEQUENCE  
75  
76 FEATURE:  
77 NAME/KEY: misc feature  
78 LOCATION: 13723..14319  
79 OTHER INFORMATION: /label= ORF13SUBSEQUENCE  
80  
81 FEATURE:  
82 NAME/KEY: misc feature  
83 LOCATION: 15655..16210  
84 OTHER INFORMATION: /label= ORF14SUBSEQUENCE  
85  
86 FEATURE:  
87 NAME/KEY: misc feature  
88 LOCATION: complement (16517..17545)

1 OTHER INFORMATION: /label= ORF15SUBSEQUENCE  
2  
3 FEATURE:  
4 NAME/KEY: misc feature  
5 LOCATION: complement (17737..18189)  
6  
7 OTHER INFORMATION: /label= ORF16SUBSEQUENCE  
8  
9 FEATURE:  
10 NAME/KEY: misc feature  
11 LOCATION: complement (18177..18743)  
12  
13 OTHER INFORMATION: /label= ORF17SUBSEQUENCE  
14  
15 FEATURE:  
16 NAME/KEY: misc feature  
17 LOCATION: complement (19031..19390)  
18  
19 OTHER INFORMATION: /label= ORF18SUBSEQUENCE  
20  
21 US-08-458-831-19  
22  
23 Query Match 6.8%; Score 121.6; DB 1; Length 21126;  
24 Best Local Similarity 45.7%; Pred. No. 8.4e-32;  
25 Matches 540; Conservative 0; Mismatches 634; Indels 8; Gaps 3;  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022

Db 8052 CTTTCATGTTGGGGTTCAACTTTTCAGGCTCTGATGCGCATTAATAATTCAGAAA 8111  
Qy 1456 GTGCTGTGAGCCGACCATATGCTCATGTTCAAGTTCAGGGGATTCAGAAAGAAAACA 1515  
Db 8112 AGCCATCGAAAGACCACTCTGTTTGTATCCGTGCTGAGTGAAGAGAGGGCGG 8171  
Qy 1516 AAAATAAAGATTAAGCTTAAGAGCGGATATCTGAACCTTAATGATTAAGTGTGTCACA 1575  
Db 8172 AGATTGAAGGATCTTGAAACACGATTCGCGGTGTTTTCAGAGTCATCATTTGGC 8231  
Qy 1576 TCTGAGATCCGAATATTCACACTAGGCAATTCCTGACATGCGATACCAATATTTTCAG 1635  
Db 8232 GCGAGTCTAGGCGCGCTAGACATGATACAGACTGCGCGGATGAGACTTCTTCAGC 8291  
Qy 1636 GCACCATGTAACCAAGGGGTTGATACAGCATATGACAGATGCTGAACCTTCTCTG 1695  
Db 8292 TACATATCGAACCCGCGTGGAACTCGTCTGCGCTGCAATTAGCATCTTTCATG 8351  
Qy 1696 ATGACTGAACGAAATCTGTTAGACCATATCTCCGCTTGTGCTCATGAGCGG 1755  
Db 8352 GTACGAAAGAAAAGTTTGGGTAACTCCGGCATCCAGAGATGATAGACCGATGG 8411  
Qy 1756 ATGCAAAAGACAGTGTATGCTGCTGACTATGATGTCGAGAT 1797  
Db 8412 CTGTCCGTAGCTGTGTCATGACATCGAATCGCAGCT 8453

## RESULT 9

US-09-230-388-2  
; Sequence 2, Application US/09230388  
; Patent No. 6291644  
; GENERAL INFORMATION:  
; APPLICANT: Iwamoto, Mitsunori; Jung, Sang-Kee  
; TITLE OF INVENTION: NOVEL PROTEIN, ITS GENE, REAGENTS FOR INDUCING  
; FILE REFERENCE: 4703/0F214  
; CURRENT APPLICATION NUMBER: US/09/230,388  
; CURRENT FILING DATE: 1999-01-22  
; PRIOR APPLICATION NUMBER: PCT/J998/02261  
; PRIOR FILING DATE: 1998-05-22  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1575  
; TYPE: DNA  
; ORGANISM: Scomber japonicus  
US-09-230-388-2

Query Match 2.1%; Score 38.6; DB 3; Length 1575;  
Best Local Similarity 66.1%; Pred. No. 0.0092;  
Matches 72; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

Qy 713 AAGTGGCGGTCAATGCGGCTGCGCATTTCCGACTCGTGTGGCAAGCAAGTCTTCATG 772  
Db 182 ATGTGTTATAGTCGAGCTGCGATGCGGACGAGCGGCGCAAGTTACTCGAAGAG 241  
Qy 773 CTGCGGAGACGATTTACATATATGAAAGCAAGTATCTGTTGGAGG 821  
Db 242 CAGG---ACACACGATAACCATATGTGAGGCTATGATCTGTTGGAGG 287

## RESULT 10

US-09-103-840A-2  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 2.1%; Score 37; DB 3; Length 4403765;  
Best Local Similarity 57.3%; Pred. No. 16;  
Matches 67; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 709 CCGAAGTGGCGGTCAATGCGGCTGCGCATTTCCGAGCTCGTGTGGCAAGCAAGTCTT 768  
Db 993867 CCGACGTTGCCGTGTCGAGCGGTAATGTCGGAATGTGCTCGCAATTACGTTGCTG 993926  
Qy 769 CATGCTGGGAGACGATGTTACCAATATATGAAAGCAAGTATCTGTTGGAGGCAAG 825  
Db 993927 AGCGCAGGATTAATGATGTCGATCTATGAAAGGCCGACGATGTTGGCGGAAG 993983

## RESULT 11

US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
US-09-103-840A-1

Query Match 2.1%; Score 37; DB 3; Length 4411529;  
Best Local Similarity 57.3%; Pred. No. 16;  
Matches 67; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 709 CCGAAGTGGCGGTCAATGCGGCTGCGCATTTCCGAGCTCGTGTGGCAAGCAAGTCTT 768  
Db 993866 CCGACGTTGCCGTGTCGAGCGGTAATGTCGGAATGTGCTCGCAATTACGTTGCTG 993925  
Qy 769 CATGCTGGGAGACGATGTTACCAATATATGAAAGCAAGTATCTGTTGGAGGCAAG 825  
Db 993926 AGCGCAGGATTAATGATGTCGATCTATGAAAGGCCGACGATGTTGGCGGAAG 993982

## RESULT 12

US-09-415-277C-4/c  
; Sequence 4, Application US/09415277C  
; Patent No. 6531308  
; GENERAL INFORMATION:  
; APPLICANT: Hersberger, Charles  
; APPLICANT: Payson, Robert  
; TITLE OF INVENTION: Ketoreductase Gene and Protein from Yeast  
; FILE REFERENCE: X-11325A  
; CURRENT APPLICATION NUMBER: US/09/415,277C

;; CURRENT FILING DATE: 1999-10-08  
;; PRIOR APPLICATION NUMBER: US 09/182,985  
;; PRIOR FILING DATE: 1998-10-30  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 4  
;; LENGTH: 1032  
;; TYPE: DNA  
;; ORGANISM: s. cerevisiae  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)..(1032)  
;; OTHER INFORMATION:  
US-09-415-277C-4

Query Match 2.0%; Score 35.2; DB 4; Length 1032;

Best Local Similarity 53.7%; Pred. No. 0.11;

Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1498 ATTCAGAAAGGAAAGACAAATAAAGATTAAGGCTTAAGCGGATATCTGAATTAT 1557  
DB 624 ATCGAAAAGTTGAGACCAAAAGAGAGAGAGGTTGACTGTTGTTGATTTGAT 565  
QY 1558 GATTAAGGTGTTGTCATCTGCACTGCAATATCCAACTCAGCGCATTCCTGACATGC 1617  
DB 564 GTGATCTTCATCTCTTTGTGAACCTCCAGCGACCTTTTCAGCAAACTTCTTGATGC 505  
QY 1618 GATACCAATATTTTTC 1633  
DB 504 AAAGTAGCATTTATC 489

## RESULT 13

US-09-415-277C-6/c  
;; Sequence 6, Application US/09415277C  
;; Patent No. 653308  
;; GENERAL INFORMATION:  
;; APPLICANT: Hershenberger, Charles  
;; TITLE OF INVENTION: Ketoreductase Gene and Protein from Yeast  
;; FILE REFERENCE: X-11325A  
;; CURRENT APPLICATION NUMBER: US/09/415, 277C  
;; CURRENT FILING DATE: 1999-10-08  
;; PRIOR APPLICATION NUMBER: US 09/182,985  
;; PRIOR FILING DATE: 1998-10-30  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 6  
;; LENGTH: 1032  
;; TYPE: RNA  
;; ORGANISM: s. cerevisiae  
US-09-415-277C-6

Query Match 2.0%; Score 35.2; DB 4; Length 1032;

Best Local Similarity 53.7%; Pred. No. 0.11;

Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1498 ATTCAGAAAGGAAAGACAAATAAAGATTAAGGCTTAAGCGGATATCTGAATTAT 1557  
DB 624 ATCGAAAAGTTGAGACCAAAAGAGAGAGAGGTTGACTGTTGTTGATTTGAT 565  
QY 1558 GATTAAGGTGTTGTCATCTGCACTGCAATATCCAACTCAGCGCATTCCTGACATGC 1617  
DB 564 GTGATCTTCATCTCTTTGTGAACCTCCAGCGACCTTTTCAGCAAACTTCTTGATGC 505  
QY 1618 GATACCAATATTTTTC 1633  
DB 504 AAAGTAGCATTTATC 489

;; Patent No. 6387658  
;; GENERAL INFORMATION:  
;; APPLICANT: Luo, Ying  
;; APPLICANT: Huang, Betty  
;; TITLE OF INVENTION: No. 6387658e1 PCNA-Associated Cell Cycle Proteins, Compositions  
;; FILE REFERENCE: A68291/DJB/RMS/DAV  
;; CURRENT APPLICATION NUMBER: US/09/404,627  
;; CURRENT FILING DATE: 1999-09-23  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 3  
;; LENGTH: 3593  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-404-627-3

Query Match 2.0%; Score 35.2; DB 4; Length 3593;

Best Local Similarity 57.1%; Pred. No. 0.29;

Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1503 GAAGGAAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTATGATA 1562  
DB 2397 GAGAAAATAATTAAATACAGATTAACATTCATTTCACAGATGTCACAGATCAGAA 2456  
QY 1563 GGTGTTGTCATCTGCACTGCAATATCCAACTCAGCGCATTCCTGACATGC 1614  
DB 2457 GATAGTAGTCTCATCTGTCTCTACCAATACACAACTGAGAGATTCCTCCGACA 2508

## RESULT 15

US-09-404-627-1  
;; Sequence 1, Application US/09404627  
;; Patent No. 6387658  
;; GENERAL INFORMATION:  
;; APPLICANT: Luo, Ying  
;; APPLICANT: Huang, Betty  
;; TITLE OF INVENTION: No. 6387658e1 PCNA-Associated Cell Cycle Proteins, Compositions  
;; FILE REFERENCE: A68291/DJB/RMS/DAV  
;; CURRENT APPLICATION NUMBER: US/09/404,627  
;; CURRENT FILING DATE: 1999-09-23  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 1  
;; LENGTH: 4205  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-404-627-1

Query Match 2.0%; Score 35.2; DB 4; Length 4205;

Best Local Similarity 57.1%; Pred. No. 0.33;

Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1503 GAAGGAAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAATTATGATA 1562  
DB 3009 GAGAAAATAATTAAATACAGATTAACATTCATTTCACAGATGTCACAGATCAGAA 3068  
QY 1563 GGTGTTGTCATCTGCACTGCAATATCCAACTCAGCGCATTCCTGACATGC 1614  
DB 3069 GATAGTAGTCTCATCTGTCTCTACCAATACACAACTGAGAGATTCCTCCGACA 3120

Search completed: November 23, 2003, 21:27:11  
Job time: 119.342 secs

RESULT 14  
US-09-404-627-3  
; Sequence 3, Application US/09404627

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 19:44:15 : Search time 557.252 Seconds  
(without alignments)  
10562.608 Million cell updates/sec

Title: US-09-434-837-10\_COPY\_1\_1801

Perfect score: 1801  
Sequence: 1 atgcagcttcaccctctcct.....catgagtcgcagatccga 1801

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2163961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/prodata/1/pubpna/US09C\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1801	100.0	2268	14	US-10-187-339-10
2	1788	99.3	1807	14	US-10-187-339-7
3	1788	99.3	2544	14	US-10-187-339-9
4	40.8	2.3	1049	12	US-10-140-472-358
5	40.8	2.3	1049	12	US-10-141-761-358
6	40.8	2.3	1049	12	US-10-142-885-358
7	40.8	2.3	1049	12	US-10-158-790-358
8	40.8	2.3	1049	12	US-10-137-871-358
9	40.8	2.3	1049	12	US-10-140-805-358
10	40.8	2.3	1049	12	US-10-140-864-358
11	40.8	2.3	1049	12	US-10-140-923-358
12	40.8	2.3	1049	12	US-10-141-756-358
13	40.8	2.3	1049	12	US-10-141-759-358
14	40.8	2.3	1049	14	US-10-123-155-358
15	40.8	2.3	1049	15	US-10-146-731-358
16	38.6	2.1	1575	9	US-09-912-176-2

17	37	2.1	1007	12	US-10-027-632-10540	Sequence 10540, A
18	37	2.1	1007	13	US-10-027-632-10540	Sequence 10540, A
19	37	2.1	161562	15	US-10-081-327-40	Sequence 40, Appl
20	36.2	2.0	3758	11	US-09-948-820-42	Sequence 42, Appl
21	36	2.0	671	14	US-10-184-644-346	Sequence 346, App
22	36	2.0	671	14	US-10-184-644-346	Sequence 346, App
23	35.6	2.0	3055	12	US-10-027-632-177179	Sequence 177179, App
24	35.6	2.0	3055	13	US-10-027-632-177179	Sequence 177179, App
25	35.2	2.0	1032	9	US-09-415-2778-4	Sequence 4, Appl1
26	35.2	2.0	1032	9	US-09-415-2778-4	Sequence 4, Appl1
27	35	1.9	368004	10	US-09-949-654-3	Sequence 3, Appl1
28	34.4	1.9	2180	14	US-10-168-274-30	Sequence 30, Appl
29	34.4	1.9	2188	14	US-10-103-313-86	Sequence 86, Appl
30	34.4	1.9	2192	12	US-10-240-965-134	Sequence 134, App
31	34.4	1.9	2281	12	US-10-320-351-16	Sequence 16, Appl
32	34.4	1.9	2291	12	US-10-320-351-16	Sequence 17, Appl
33	34.4	1.9	11745	12	US-10-240-453-205	Sequence 205, App
34	34	1.9	585	12	US-10-032-585-6692	Sequence 6692, App
35	34	1.9	585	13	US-10-027-632-217634	Sequence 217634, App
36	34	1.9	585	13	US-10-027-632-217634	Sequence 217634, App
37	34	1.9	1572	14	US-10-230-026-25	Sequence 29, Appl
38	33.8	1.9	1302	9	US-09-815-242-9647	Sequence 9647, App
39	33.6	1.9	413	14	US-10-184-644-156	Sequence 156, App
40	33.6	1.9	413	14	US-10-184-644-156	Sequence 156, App
41	33.6	1.9	1094	14	US-10-103-313-225	Sequence 225, App
42	33.4	1.9	708	12	US-10-140-472-298	Sequence 298, App
43	33.4	1.9	708	12	US-10-141-761-298	Sequence 298, App
44	33.4	1.9	708	12	US-10-142-885-298	Sequence 298, App
45	33.4	1.9	708	12	US-10-158-790-298	Sequence 298, App

#### ALIGNMENTS

RESULT 1  
US-10-187-339-10  
; Sequence 10, Application US/10187339  
; Publication No. US20030084478A1  
; GENERAL INFORMATION:  
; APPLICANT: Ream, Walt et al.,  
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
; TITLE OF INVENTION: Methods and Compositions for Producing the Same  
; FILE REFERENCE: 53629  
; CURRENT APPLICATION NUMBER: US/10/187,339  
; PRIOR FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US/09/434,837  
; PRIOR FILING DATE: 1999-11-04  
; PRIOR APPLICATION NUMBER: 60/107,185  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 2268  
; TYPE: DNA  
; ORGANISM: Agrobacterium tumefaciens  
US-10-187-339-10

Query Match 100.0%; Score 1801; DB 14; Length 2268;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGTGAGCTTACCTCTCTTGATACCGATGATCTCCCAACCAAAATGTGGAT	60
DB	1	ATGTGAGCTTACCTCTCTTGATACCGATGATCTCCCAACCAAAATGTGGAT	60
QY	61	CTGACATGATGATGATTAAGCGATGATGATGATGATGATGATGATGATGAT	120
DB	61	CTGACATGATGATGATTAAGCGATGATGATGATGATGATGATGATGATGAT	120
QY	121	CGAGAACTTCTTACGAGGAGAGAGATTAAGCTTCCAGGATGAGCGTGGGTTA	180
DB	121	CGAGAACTTCTTACGAGGAGAGAGATTAAGCTTCCAGGATGAGCGTGGGTTA	180

181 GCTTGAAGAGGCTGCGGATGCTCGCTCCCGAGATCTAGCTGGTGAAGATGCA 240  
181 GCTTGAAGAGGCTGCGGATGCTCGCTCCCGAGATCTAGCTGGTGAAGATGCA 240  
241 GTTCTCTCCGCTTATATATATATATATATATATATATATATATATATATAT 300  
241 GTTCTCTCCGCTTATATATATATATATATATATATATATATATATATATAT 300  
301 CTTTGGGCGGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
301 CTTTGGGCGGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
361 GATTTCCTCCAGAGCAACTATATCAAGCCCTGTTTGTGAGAGAGAGAGAGAG 420  
361 GATTTCCTCCAGAGCAACTATATCAAGCCCTGTTTGTGAGAGAGAGAGAGAG 420  
421 CCGATGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATGATCTT 480  
421 CCGATGATCTTATGATCTTATGATCTTATGATCTTATGATCTTATGATCTT 480  
481 CTGCGCATGCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
481 CTGCGCATGCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
541 CTTGAAGAGGCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
541 CTTGAAGAGGCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
601 TCGGAGAGGCTCTTCCCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
601 TCGGAGAGGCTCTTCCCAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
661 TGTTCGATAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
661 TGTTCGATAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
721 GTCAATGCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
721 GTCAATGCGCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
781 GAGCATTTACAT 840  
781 GAGCATTTACAT 840  
841 TTTAAGAGAGGCTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
841 TTTAAGAGAGGCTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
901 TTTGCTGCTGCTTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
901 TTTGCTGCTGCTTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
961 CCGGAG 1020  
961 CCGGAG 1020  
1021 CAGTGCACAG 1080  
1021 CAGTGCACAG 1080  
1081 GGTTCATGAG 1140  
1081 GGTTCATGAG 1140  
1141 TCAAG 1200  
1141 TCAAG 1200  
1201 GAGTGCCT 1260  
1201 GAGTGCCT 1260  
1261 CATGAG 1320

1261 CATGAG 1320  
1321 TTTGCTGCTGCTTCT 1380  
1321 TTTGCTGCTGCTTCT 1380  
1381 GAG 1440  
1381 GAG 1440  
1441 GAGTGCCT 1500  
1441 GAGTGCCT 1500  
1501 CAG 1560  
1501 CAG 1560  
1561 AAGTGCCT 1620  
1561 AAGTGCCT 1620  
1621 ACCAATATTTTTCAG 1680  
1621 ACCAATATTTTTCAG 1680  
1681 TCAAACT 1740  
1681 TCAAACT 1740  
1741 GTCTCTATGAG 1800  
1741 GTCTCTATGAG 1800  
1801 A 1801  
1801 A 1801

RESULT 2  
US-10-187-339-7  
; Sequence 7, Application US/10187339  
; Publication No. US20030084478A1  
; GENERAL INFORMATION:  
; APPLICANT: Ream, Malt et al.  
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
; FILE REFERENCE: Methods and Compositions for Producing the Same  
; CURRENT APPLICATION NUMBER: US/10/187,339  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US/09/434,837  
; PRIOR FILING DATE: 1999-11-04  
; PRIOR APPLICATION NUMBER: 60/107,185  
; PRIOR FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 1807  
; TYPE: DNA  
; ORGANISM: Agrobacterium tumefaciens  
US-10-187-339-7

Query Match 99.3%; Score 1788; DB 14; Length 1807;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12 ACCTCTCTCTGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 71  
18 ACCTCTCTCTGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77  
72 CGATTAAG 131

Db 78 CGATTAAGCCGATGAATTGGAACCGCAGGTTTCCGATGCTTCTTAGAACGAGAACCTTC 137  
Qy 132 TAGGGGAGGAGGATTAATCTCAAGCTCCACGAGTGGAGGCTGGGTAGCTTGCAAAAG 191  
Db 138 TAGGGGAGGAGGATTAATCTCAAGCTCCACGAGTGGAGGCTGGGTAGCTTGCAAAAG 197  
Qy 192 GCTGGCCGATGTCGCTTCCCGAGATCTGAGCTGTGGAAGGATGACAGTCTCTCCG 251  
Db 198 GCTGGCCGATGTCGCTTCCCGAGATCTGAGCTGTGGAAGGATGACAGTCTCTCCG 257  
Qy 252 TTATATCTAATATGGAAGAAATTTCTGGGGCGGATCTTGAAATCGAAACCTTGGGCG 311  
Db 258 TTATATCTAATATGGAAGAAATTTCTGGGGCGGATCTTGAAATCGAAACCTTGGGCG 317  
Qy 312 GGCAACGATGATGTCGCTGTCGATCGACTTGGAACCACTTGATGATTTCTCCGA 371  
Db 318 GGCAACGATGATGTCGCTGTCGATCGACTTGGAACCACTTGATGATTTCTCCGA 377  
Qy 372 AGCAACAATATCCAAAGCCCTGTTTGTGAGCGGTAAGAAAGATGTCACCGATGATCT 431  
Db 378 AGCAACAATATCCAAAGCCCTGTTTGTGAGCGGTAAGAAAGATGTCACCGATGATCT 437  
Qy 432 TAGCTATTTGTCGCTCAATTTCTTAAGACTGCGGCTTTGAAACCTTCCCATGCC 491  
Db 438 TAGCTATTTGTCGCTCAATTTCTTAAGACTGCGGCTTTGAAACCTTCCCATGCC 497  
Qy 492 GCTGTAAGAAATGGAAGCAATGCAATGCTGACCGGTTTACCAATACCTTTGAAGGCG 551  
Db 498 GCTGTAAGAAATGGAAGCAATGCAATGCTGACCGGTTTACCAATACCTTTGAAGGCG 557  
Qy 552 CGTGCCATTTGACATGTCGCTTATGTCGAAACCTGATGTCGAAAGGTTCCGACAGTTC 611  
Db 558 CGTGCCATTTGACATGTCGCTTATGTCGAAACCTGATGTCGAAAGGTTCCGACAGTTC 617  
Qy 612 CTTTCCAAATGCAATGCTGCTTAAGACTGCAAGCCGTTTGTGCAACATGTTCCGATG 671  
Db 618 CTTTCCAAATGCAATGCTGCTTAAGACTGCAAGCCGTTTGTGCAACATGTTCCGATG 677  
Qy 672 TGAACGATGTCGCTTCTTCCGAGGATGTTCTTAAGCCGAAAGTGGCGCTTATGGCGC 731  
Db 678 TGAACGATGTCGCTTCTTCCGAGGATGTTCTTAAGCCGAAAGTGGCGCTTATGGCGC 737  
Qy 732 TGCAATTTCCGAGCTCGTGGTGGCAACGAACTGCTCATGCTGGGTTAGACATGTTAC 791  
Db 738 TGCAATTTCCGAGCTCGTGGTGGCAACGAACTGCTCATGCTGGGTTAGACATGTTAC 797  
Qy 792 AATATATGAAGCAAGTATGCTGTGGAAGCAAGCTTTGTCATGCTTTCAAGGACGC 851  
Db 798 AATATATGAAGCAAGTATGCTGTGGAAGCAAGCTTTGTCATGCTTTCAAGGACGC 857  
Qy 852 TCCTAGTGTGTCGCGCAAAATGGGGGAGATGCGATTTCTCTGCTGCTGATTTGCTTGT 911  
Db 858 TCCTAGTGTGTCGCGCAAAATGGGGGAGATGCGATTTCTCTGCTGCTGATTTGCTTGT 917  
Qy 912 TTTCTTCTCGAGGCTTACGAGCTGTTTGCATGAGGCGGTTCCCAAATCCCGACAGT 971  
Db 918 TTTCTTCTCGAGGCTTACGAGCTGTTTGCATGAGGCGGTTCCCAAATCCCGACAGT 977  
Qy 972 CGACACTTACTTGTGTACCAAGCGCTCAATACATGTGAAAGCCGGGAGCTGCCACC 1031  
Db 978 CGACACTTACTTGTGTACCAAGCGCTCAATACATGTGAAAGCCGGGAGCTGCCACC 1037  
Qy 1032 GAAGCTGTTCATGCGCTTTAACAAGGTTGGCGTGGCTTTGTAAGACGCTTTTATGA 1091  
Db 1038 GAAGCTGTTCATGCGCTTTAACAAGGTTGGCGTGGCTTTGTAAGACGCTTTTATGA 1097  
Qy 1092 GCGAGATATGTCGCTTCCGCTGCTGCTATTTACTCAGAGCTTGAAATCAGAGACAT 1151  
Db 1098 GCGAGATATGTCGCTTCCGCTGCTGCTATTTACTCAGAGCTTGAAATCAGAGACAT 1157  
Qy 1152 TAGTGGGCTCATGACTCTGCGCAAAATTTGGCTGAACCGTTTCCGAGGAGTCTTCTC 1211  
Db 1158 TAGTGGGCTCATGACTCTGCGCAAAATTTGGCTGAACCGTTTCCGAGGAGTCTTCTC 1217

Qy 1212 TTCAGGATAGAGGATCTTTCTGGGCAACATCTCTCCGNGGTAAACATGAGTTTT 1271  
Db 1218 TTCAGGATAGAGGATCTTTCTGGGCAACATCTCTCTGGGTAAACATGAGTTTT 1277  
Qy 1272 CCTCATGATTTGGGACCTTATTCAGCTAATGGAATAGATCTGGCGGGTTGGTCAGT 1331  
Db 1278 CCTCATGATTTGGGACCTTATTCAGCTAATGGAATAGATCTGGCGGGTTGGTCAGT 1337  
Qy 1332 TTTTGAAGCGGTTTATGAGATCTCTCGCTTGGTCATCAACGATATGAAGAAATCA 1391  
Db 1338 TTTTGAAGCGGTTTATGAGATCTCTCGCTTGGTCATCAACGATATGAAGAAATCA 1397  
Qy 1392 GCGGATGTCGCTTGAAGGATCTCAGAACTTCAGCTCGGATGATGATGATTTAA 1451  
Db 1398 GCGGATGTCGCTTGAAGGATCTCAGAACTTCAGCTCGGATGATGATGATTTAA 1457  
Qy 1452 CGGTGTGTCGTGAGCGCAGCATATGCGATGTTCAAGTCAGGCGATTCAAGAGAAA 1511  
Db 1458 CGGTGTGTCGTGAGCGCAGCATATGCGATGTTCAAGTCAGGCGATTCAAGAGAAA 1517  
Qy 1512 GACAAAATTAAGATTAAGCTTAAGCGGATATTTGAACCTTTATGATTAAGTGTGT 1571  
Db 1518 GACAAAATTAAGATTAAGCTTAAGCGGATATTTGAACCTTTATGATTAAGTGTGT 1577  
Qy 1572 CACATCTGACTGCAATATCAACTCAGGATGTCGCAATGCGATACCAATTTTT 1631  
Db 1578 CACATCTGACTGCAATATCAACTCAGGATGTCGCAATGCGATACCAATTTTT 1637  
Qy 1632 TCAGGACCAAGTAACCAAGCGGTTGATTAACAGCCATATGACAGATTCGTCAAACTTT 1691  
Db 1638 TCAGGACCAAGTAACCAAGCGGTTGATTAACAGCCATATGACAGATTCGTCAAACTTT 1697  
Qy 1692 CCGTATGACTGAACGAAATTTCTGTTAAGCAATATCTCCGCTTGTGTCTCATGGA 1751  
Db 1698 CCGTATGACTGAACGAAATTTCTGTTAAGCAATATCTCCGCTTGTGTCTCATGGA 1757  
Qy 1752 CCGGATCGCAAAAGCATATTTCCCTGACATATGATGTCGCGAGATCC 1799  
Db 1758 CCGGATCGCAAAAGCATATTTCCCTGACATATGATGTCGCGAGATCC 1805

RESULT 3  
US-10-187-339-9  
; Sequence 9, Application US/10187339  
; Publication No. US20030084478A1  
; GENERAL INFORMATION:  
; APPLICANT: Ream, Walt et al.,  
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and  
; FILE OF INVENTION: Methods and Compositions for Producing the Same  
; FILE REFERENCE: 53629  
; CURRENT APPLICATION NUMBER: US/10/187,339  
; PRIOR FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US/09/434,837  
; PRIOR FILING DATE: 1999-11-04  
; PRIOR FILING DATE: 1998-11-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 2544  
; TYPE: DNA  
; ORGANISM: Agrobacterium tumefaciens  
US-10-187-339-9

Query Match 99.3%; Score 1788; DB 14; Length 2544;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 12 ACCTCTCTTGAATTAACGATGCGATCATCTCCCAACCAAAATGATGATCTGACAAATGAT 71  
Db 318 ACCTCTCTTGAATTAACGATGCGATCATCTCCCAACCAAAATGATGATCTGACAAATGAT 377

QY 72 CGATTAAGCGGATGATTAATGGACCGCAGGGTTTCGATGCTCTTGAAGCAGAGGCTTC 131  
DB 378 CGATTAAGCGGATGATTAATGGACCGCAGGGTTTCGATGCTCTTGAAGCAGAGGCTTC 437  
QY 132 TAGGGGAGGAGGATTAATCTCAAGCTTCCAGAGTGCAGCGCTGGGTTAGCTTGGCAAG 191  
DB 438 TAGGGGAGGAGGATTAATCTCAAGCTTCCAGAGTGCAGCGCTGGGTTAGCTTGGCAAG 497  
QY 192 GCTGGCCGATGCTGCTTCCCGAGATCTCAGCTGGTGGAAAGGTACAGTTCCTCCGC 251  
DB 498 GCTGGCCGATGCTGCTTCCCGAGATCTCAGCTGGTGGAAAGGTACAGTTCCTCCGC 557  
QY 252 TTATATCTATTTGGCAAGAAATTTCTGGGGCGGATCTTGAATCGAAACCTTGGGGCG 311  
DB 558 TTATATCTATTTGGCAAGAAATTTCTGGGGCGGATCTTGAATCGAAACCTTGGGGCG 617  
QY 312 GGCACAGTAGTGTCTGCTTGCATCGACTTGGACACCATTTGACATTTCTCGA 371  
DB 618 GGCACAGTAGTGTCTGCTTGCATCGACTTGGACACCATTTGACATTTCTCGA 677  
QY 372 AGCACAATAATCCAAAGCCCTGTTTTCGAGAGGGGTAAGATGTCACCGATGATCT 431  
DB 678 AGCACAATAATCCAAAGCCCTGTTTTCGAGAGGGGTAAGATGTCACCGATGATCT 737  
QY 432 TAGCATTTTGTGGCCATTTCAATCTCTAAGACTGCGGCTTTGGAACCTTGGCAATGCC 491  
DB 738 TAGCATTTTGTGGCCATTTCAATCTCTAAGACTGCGGCTTTGGAACCTTGGCAATGCC 797  
QY 492 GCTGTACGAGATGCGACGATGAATGCGTTACCGGGTTTACATTAACCTTGAAGGGGC 551  
DB 798 GCTGTACGAGATGCGACGATGAATGCGTTACCGGGTTTACATTAACCTTGAAGGGGC 857  
QY 552 CGTGGCCATTTGACATGCTTACCTTAATGTCGAAACCTGATGAGGTTCCGAGGTTTC 611  
DB 858 CGTGGCCATTTGACATGCTTACCTTAATGTCGAAACCTGATGAGGTTCCGAGGTTTC 917  
QY 612 CTCTTCAACATCGACTTGTCTACGACTGACAGCCGTTTGTACCAATGTTCCGATAG 671  
DB 918 CTCTTCAACATCGACTTGTCTACGACTGACAGCCGTTTGTACCAATGTTCCGATAG 977  
QY 672 TGGACGAGTGGCTTCTTTCCGAGAGTGTCTTAACCGGAAGTGGCGCTCATTTGCGC 731  
DB 978 TGGACGAGTGGCTTCTTTCCGAGAGTGTCTTAACCGGAAGTGGCGCTCATTTGCGC 1037  
QY 732 TGGCATTTCCGAGACTGCTGGTGGCAACGAACTGCTTCACTGGGGTAAACGATGTTAC 791  
DB 1038 TGGCATTTCCGAGACTGCTGGTGGCAACGAACTGCTTCACTGGGGTAAACGATGTTAC 1097  
QY 792 AATATATGAAGCAAGTATGCTTTGGAGGCAAGCTTGGTCAATGCTTTCAAGGACGC 851  
DB 1098 AATATATGAAGCAAGTATGCTTTGGAGGCAAGCTTGGTCAATGCTTTCAAGGACGC 1157  
QY 852 TCCTAGTGTGCTGGCCGAAATGGGGGCGATGCGATTTCTCTGCTGCTCATTTCTGCTGT 911  
DB 1158 TCCTAGTGTGCTGGCCGAAATGGGGGCGATGCGATTTCTCTGCTGCTCATTTCTGCTGT 1217  
QY 912 TTTCTTCTGAGAGGCTTACGCGCTGTCTTGAATAGAGCCGTTCCCAATCCCGCAAGT 971  
DB 1218 TTTCTTCTGAGAGGCTTACGCGCTGTCTTGAATAGAGCCGTTCCCAATCCCGCAAGT 1277  
QY 972 CGACACTTAATCTGCTTACCAAGGCGTCAATATACATGTAAGAAACCGGGACGCTGCCACC 1031  
DB 1278 CGACACTTAATCTGCTTACCAAGGCGTCAATATACATGTAAGAAACCGGGACGCTGCCACC 1337  
QY 1032 GAAGCTTTCATCGCGTTTACAAACGCTTGGCGTCTTCTTGAAGACGCTTTTCAATGA 1091  
DB 1338 GAAGCTTTCATCGCGTTTACAAACGCTTGGCGTCTTCTTGAAGACGCTTTTCAATGA 1397  
QY 1092 GCGAGATATTTGCTTGGCTTGCCTGCTGCTATTAATCAAGGCTTGAATATAGACAT 1151  
DB 1398 GCGAGATATTTGCTTGGCTTGCCTGCTGCTATTAATCAAGGCTTGAATATAGACAT 1457  
QY 1152 TAGGTGGGCTCATGATCTCTGGCAAAATTTGGCTGAACGCTTTCCGAGGAGTCTCTC 1211

DB 1458 TAGGTGGGCTCATGATCTCTGGCAAAATTTGGCTGAACCGTTTGGGAGGGAGTCTCTC 1517  
QY 1212 TTCAAGGATTAAGAGGATCTTTCTGGGACACATCTCTCTGGTGTAAACATGAGTTT 1271  
DB 1518 TTCAAGGATTAAGAGGATCTTTCTGGGACACATCTCTCTGGTGTAAACATGAGTTT 1577  
QY 1272 CCTCATGATTTGGAGCCATTAATCAAGCTAATGGAATAGATCTGGCGGGTTTGGTCCAGT 1331  
DB 1578 CCTCATGATTTGGAGCCATTAATCAAGCTAATGGAATAGATCTGGCGGGTTTGGTCCAGT 1637  
QY 1332 TTTTGAAGCGGGTTTATTAAGATCTCTCGCTTGTATCAACGATATGAAGAAATCA 1391  
DB 1638 TTTTGAAGCGGGTTTATTAAGATCTCTCGCTTGTATCAACGATATGAAGAAATCA 1697  
QY 1392 GCGGATGTGCTTGAAGGAATCTCAAGCTTCAAGCTCGGATGCTAATGTTGTTAA 1451  
DB 1698 GCGGATGTGCTTGAAGGAATCTCAAGCTTCAAGCTCGGATGCTAATGTTGTTAA 1757  
QY 1452 CGGTGTGCTTGAAGCGAGCATATGCCATGTTCAAGTCAGGGCGATTGAGAGAA 1511  
DB 1758 CGGTGTGCTTGAAGCGAGCATATGCCATGTTCAAGTCAGGGCGATTGAGAGAA 1817  
QY 1512 GACAAAATTAAGATTAAGGCTTAAGACCGGATATCTGAACCTTATGATTAAGTGTGT 1571  
DB 1818 GACAAAATTAAGATTAAGGCTTAAGACCGGATATCTGAACCTTATGATTAAGTGTGT 1877  
QY 1572 CACATCTGACCTGGCAATATCCAACTCAGGCTTGGCTGACATGCGATCCAAATATTT 1631  
DB 1878 CACATCTGACCTGGCAATATCCAACTCAGGCTTGGCTGACATGCGATCCAAATATTT 1937  
QY 1632 TCAGGACAGCATGAACCAAGCGGTTGATTAACAGCCATATGACAGATCTCAAACTCT 1691  
DB 1938 TCAGGACAGCATGAACCAAGCGGTTGATTAACAGCCATATGACAGATCTCAAACTCT 1997  
QY 1692 CTTGATGACTGAACGAAATTTCTGTTAGACATATCTCTCCGTCTTGTCTCATGA 1751  
DB 1998 CTTGATGACTGAACGAAATTTCTGTTAGACATATCTCTCCGTCTTGTCTCATGA 2057  
QY 1752 CGGATTCGCAAAAGCACTGATTTGCTGGAATGATGATGAGTCCAGAGATTC 1799  
DB 2058 CGGATTCGCAAAAGCACTGATTTGCTGGAATGATGATGAGTCCAGAGATTC 2105

RESULT 4  
US-10-140-472-358  
; Sequence 358, Application US/10140472  
; Publication No. US2003013888A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Geriltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C168  
; CURRENT APPLICATION NUMBER: US/10/140,472  
; PRIOR FILING DATE: 2002-05-06  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550

SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-140-472-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;  
Best Local Similarity 5.4%; Pred. No. 0.016;  
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCGCCGCTTGAACATGGTATGCTGAACCTGAGCTGAAG 597  
DB 298 HSNLSLQHPVPRFKNINKLOELDSQNFLEKIDGAKFLHFLPSLIQLDLSFNFELQVVR 357  
QY 598 GGTTCGAGAGTTCCTTCCACATGACTGCTGACAGCTGACAGCCGTTTGGAC 657  
DB 358 ASNMLSQAFSSLSKLRLRGVFEKLSFNLSPLHNLQNLVDLGTNFIKIANLSMF 417  
QY 658 CAATGTCGATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCCGAAGCATG 717  
DB 418 KQFKRLKVIDLSVKNISPSGDSSEVFCSNARTSVSEYEPVLEQLHYFRYDKVARSCHR 477  
QY 718 GCGGTATTCGCGCTGCGCATTTCCGAGCTGCTGTCGCAAGCAACTGCTTACGCTGG 777  
DB 478 KNEASFMSVNSCYKYGQTLDSKNSIFPVKSSDFQLFLKCLNLSGNLISQTLNGSE 537  
QY 778 GTAGACATGTTACATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTGTACAT 837  
DB 538 FQPLAELRYLDFSNRDLHSTAFEBLHLEVLIDSSNSHYFQSGITTHMLNFTNKLKV 597  
QY 838 GCTTTCGAGGAGCTCCTTGAATGTCGTCGCAAGATGGGGCGATTCGATTCCTCTGCT 897  
DB 598 LQKLMMNDNDISSSTRWESSELRLEFRGNHLDVLMREGDNRVYLQFLNKLKLELDI 657  
QY 898 GCATTCGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTTCTGATGAGCGCGTCCCA 957  
DB 658 SKNSLSFLPSGVFDGMPNKLNLKSLKGLKFSWKLOCLKNLETLIDLSHNOITTVPER 717  
QY 958 AATCCGCGACAGTGCAGCACTTACTGCTGTACCAAGCGCTCCCATATGATGTGAAGCC 1017  
DB 718 LSNCSLSLKNLILKNNQIRSLTYKFLQDAFQRLYLDLSSNKMIOIKTSFPEVNLNKLK 777  
QY 1018 GGGCAGCTGCCACGAAGCTGTTCCATCGGCTTACCAAGCTGCGCTGCTTGAAG 1077  
DB 778 LLLHNRFLCTCAVWVWVWVHTVTIPIYLAIDVTCVGPAAKGSVISLDYTCELD 837  
QY 1078 GACGCTTTCATGAGCAGATATGTTGCTGCTGCTGCTGCTATTAATCAAGCGCTTG 1137  
DB 838 TNLILFSLISVSLFLWVMTASHLYFMVWYIYHFCAKIKGYORLISPDCCYDAFIY 897  
QY 1138 AATCAGACACATTAGTGGCTCATGACTCCTCGCAATTTGGCTGAACGCTTCCGG 1197  
DB 898 DTKDPAVTEWVLAELVAKLEDPREKHFNLCLBERDMLPGQVLENLSSQISLTKTVFWM 957  
QY 1198 AGGAGTCTTCTCTTCAGGATAGAGAGATCTTTCGGAACACAT 1245  
DB 958 TDKYAKTENFKIAFYLSHQRLMDEKVDVILLIFLEKPFQKSKFLQLRK 1005

RESULT 5  
US-10-141-761-358  
Sequence 358, Application US/10141761  
Publication No. US20030148432A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Geo, Wei-Qiang  
APPLICANT: Geritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Collin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P330R1C198  
CURRENT APPLICATION NUMBER: US/10/141,761  
CURRENT FILING DATE: 2002-05-08  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-141-761-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;  
Best Local Similarity 5.4%; Pred. No. 0.016;  
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCGCGCTGCAATGATGATGCTTATGCTGAACCTGAGCTGAAG 597  
DB 298 HSNLSLQHPVPRFKNINKLOELDSQNFLEKIDGAKFLHFLPSLIQLDLSFNFELQVVR 357  
QY 598 GGTTCGAGAGTTCCTTCCACATGCACTGCTGCTGACAGCTGACAGCGTTTGGAC 657  
DB 358 ASNMLSQAFSSLSKLRLRGVFEKLSFNLSPLHNLQNLVDLGTNFIKIANLSMF 417  
QY 658 CAATGTCGATAGTGAAGCATGCGCTTCTTCCGAGAGATGCTTGAAGCAAGCATG 717  
DB 418 KQFKRLKVIDLSVKNISPSGDSSEVFCSNARTSVSEYEPVLEQLHYFRYDKVARSCHR 477  
QY 478 GCGGTATTCGCGCTGCGCATTTCCGAGCTGCTGTCGCAAGCAAGCAACTGCTTACGCTGG 777  
DB 718 LSNCSLSLKNLILKNNQIRSLTYKFLQDAFQRLYLDLSSNKMIOIKTSFPEVNLNKLK 777  
QY 778 GTAGACATGTTACATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTGTACAT 837  
DB 838 FQPLAELRYLDFSNRDLHSTAFEBLHLEVLIDSSNSHYFQSGITTHMLNFTNKLKV 597  
QY 838 GCTTTCGAGGAGCTCCTTGAATGTCGTCGCAAGATGGGGCGATTCGATTCCTCTGCT 897  
DB 598 LQKLMMNDNDISSSTRWESSELRLEFRGNHLDVLMREGDNRVYLQFLNKLKLELDI 657  
QY 898 GCATTCGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTTCTGATGAGCGCGTCCCA 957  
DB 658 SKNSLSFLPSGVFDGMPNKLNLKSLKGLKFSWKLOCLKNLETLIDLSHNOITTVPER 717  
QY 958 AATCCGCGACAGTGCAGCACTTACTGCTGTACCAAGCGCTCCCATATGATGTGAAGCC 1017  
DB 718 LSNCSLSLKNLILKNNQIRSLTYKFLQDAFQRLYLDLSSNKMIOIKTSFPEVNLNKLK 777  
QY 1018 GGGCAGCTGCCACGAAGCTGTTCCATCGGCTTACCAAGCTGCGCTGCTTGAAG 1077  
DB 778 LLLHNRFLCTCAVWVWVWVHTVTIPIYLAIDVTCVGPAAKGSVISLDYTCELD 837  
QY 1078 GACGCTTTCATGAGCAGATATGTTGCTGCTGCTGCTGCTATTAATCAAGCGCTTG 1137  
DB 838 TNLILFSLISVSLFLWVMTASHLYFMVWYIYHFCAKIKGYORLISPDCCYDAFIY 897  
QY 1138 AATCAGACACATTAGTGGCTCATGACTCCTCGCAATTTGGCTGAACGCTTCCGG 1197  
DB 898 DTKDPAVTEWVLAELVAKLEDPREKHFNLCLBERDMLPGQVLENLSSQISLTKTVFWM 957  
QY 1198 AGGAGTCTTCTCTTCAGGATAGAGAGATCTTTCGGAACACAT 1245  
DB 958 TDKYAKTENFKIAFYLSHQRLMDEKVDVILLIFLEKPFQKSKFLQLRK 1005

```
RESULT 6
US-10-142-885-358
; Sequence 358, Application US/10142885
; Publication No. US20030157604A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C248
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: US/10/142,885
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-885-358

Query Match      2.3%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.016;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCGCGCCATTGACATGATGCTTATGCTGAACCGTGAAG 597
DB 298 HNSLSLQHVPRFWKINIKQLDELDSQNFIAKEIGDAKFLHFLPSLIQDLSPFELQVYR 357
QY 598 GCTTCGGCAGGCTTCCTTCCAAACATGACATGCTGCTACACAGCAGCAGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKILIRIGYVFKELSFNLSPLHNLQNLVLDLGTNFIKIANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
DB 418 KQPRKLVLDLNVKISPSGDSSEVFCNSARTSVESYEPOVLEQLHYFRYDKYARSCR 477
QY 718 GCGGTCAATGGCGCTGACATTTCCGACACTGCTGTCGCAACGAACTGCTTACGTGG 777
DB 478 KKEASFMVSNECYKYGQTLDSKNSIFPVKSSDFQHLFLKLNLSGMLISQTLNGSE 537
QY 778 GTAGACATGTTACAATATATGAAGCAAGTATGCTTGGAGCAAGCTTGGTCACAT 837
DB 538 FQPLAEIRYLDPSNNRDLHSTAFEBELHLEVLDISNSHYFSEGITTHLNLKLV 597
QY 838 GCTTTCAGGAGCTCTAGTGTGTCGCGCAATGAGGGGCGATGCGATTTCTCTGCT 897
DB 598 LQKLMMNDNISSTSTSESLRTLEFRGNHLDVLMRSDNRYLDLFLNKLKLELDI 657
QY 898 GCATTCGCTGTTTCTTCTTCGACGCTTACGCGCTGCTTTCGATGAGCGCTTCCA 957
DB 658 SKNSLSFLPSGVFDPMPNLKNLSLAKGLKFSMKLQCLKNLETLDLSHNQLTTPER 717
QY 958 AATCCCGCACAGTCGACATTAATGCTTACACAGCGCTTCAATCATGTGAAGCC 1017
DB 718 LSNNSRLKVLKLNQIRSLTKYFLQDAFQRLDLSNNKIOMIQKTSFPEVNLNLM 777
QY 1018 GGGAGCTGCCACGAGAGCTGTTCCATCGCTTACACAGGTTGGCGTTCGTTGAAG 1077
```

```
DB 778 LLLHNRFLCTDPAVFWVMVNHTEVTIPIPLATDVTCVBPAGAKGOSVSLDLYTCBIDL 837
QY 1078 GAGGCTTTCATGAGGAGATATGCTGTGCGCTTCGCCCTGCTCTTACTACAGCGCTTG 1137
DB 838 TNLILFSLISVSLFLMVMTASHLFFWDWYVYHFCRKKIKYQRLISPDCCYDAFIYV 897
QY 1138 AATTCAGACACATTAATGATGAGGCTCATGACTCCTGCGCAATTTGGCTGAACGTTTCGG 1197
DB 898 DTGDPATVTEVLAELVAKLEDPREKHFNLCLERDMLPGQPVLENLSQSLQSKTVFWM 957
QY 1198 AGGAGCTCTTCTTCCGAGGATAGAGAGATCTTTCGCGACACAT 1245
DB 958 TDKYAKTENFKIAFYLSHORLMEKVDVILLIFLEKFPQSKRFLQRLK 1005

RESULT 7
US-10-158-790-358
; Sequence 358, Application US/10158790
; Publication No. US20030180879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C448
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US/10/158,790
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-358

Query Match      2.3%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.016;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTGAAGGGCGCGCCATTGACATGATGCTTATGCTGAACCGTGAAG 597
DB 298 HNSLSLQHVPRFWKINIKQLDELDSQNFIAKEIGDAKFLHFLPSLIQDLSPFELQVYR 357
QY 598 GCTTCGGCAGGCTTCCTTCCAAACATGACATGCTGCTTACAGCTGACAGCGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKILIRIGYVFKELSFNLSPLHNLQNLVLDLGTNFIKIANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
DB 418 KQPRKLVLDLNVKISPSGDSSEVFCNSARTSVESYEPOVLEQLHYFRYDKYARSCR 477
QY 718 GCGGTCAATGGCGCTGACATTTCCGACACTGCTGTCGCAACGAACTGCTTACGTGG 777
DB 478 KKEASFMVSNECYKYGQTLDSKNSIFPVKSSDFQHLFLKLNLSGMLISQTLNGSE 537
QY 778 GTAGACATGTTACAATATATGAAGCAAGTATGCTTGGAGCAAGCTTGGTCACAT 837
```

Db 538 FQPLAEIRYDFSNRNLIDLIHSTAEEBLHKEVLIDISSNSHYFQSEGITMLNFTKMLKV 597  
Qy 838 GCTTTCAGGAGCCTCTAGTGTGTGCGCAATGGGGCGATGCGATTTCCCTGCT 897  
Db 598 LQKLMNNDNDISSSTRTMSESLRTLEFRGNHLDVLMREGDNRVYQLFKNILKLEBDI 657  
Qy 898 GCATTTGCTTTGTTTTCTTCTCGAGCGCTTACCGCCTGTCTTGATGAGCGCGTTCCA 957  
Db 658 SKNSLFLPSGVFDGMPNMLKNLSLAKNGLKSFWSKKLOCLKNLETLIDLSHNLTVPER 717  
Qy 958 AATCCCGGCACTGACACTTACTGCTGTACCAAGCGCTCCAAATCATGTGGAAGCC 1017  
Db 718 LSNCSLSLKNLILKNNQIRSLTKYFLQDAPQLRYLIDLSNNKIQMTQKTSFPEVNLNLM 777  
Qy 1018 GGCAGCTGCGCCAGAGCTGTCATCGGCTTACCAAGCTTGGGCTGCTTGAAG 1077  
Db 778 LLLHNRFLCTCAVFWFVWVWVHTVTIPLYLATDVTCVPAHKGQSVSLDYLTBELD 837  
Qy 1078 GACGGTTTTCATGAGAGATATGTGTGCTTGGCTTGGCTGCTATTACTCAGGCTTG 1137  
Db 838 TNLILFSLISVSLFLMVMWMTASHLYFMDVWYIYHFKAKIKGYORLISPDCYDAFIY 897  
Qy 1138 AATCAGACACATTAGTGCGCTCATGACTCCTGCGAAATTTGGCTGAACGTTTGGG 1197  
Db 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLBERDMLPGQVLENLSQSIQLSKTVFWM 957  
Qy 1198 AGGAGTCTCTCTTCTCAGGATAGAGAGATCTTCTGCGCACACAT 1245  
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILILEKPFQSKFLQLRK 1005

RESULT 8  
US-10-137-871-358  
; Sequence 358, Application US/10137871  
; Publication No. US20030207350A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C153  
; CURRENT APPLICATION NUMBER: US/10/137, 871  
; PRIOR FILING DATE: 2002-05-03  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-137-871-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;  
Best Local Similarity 5.4%; Pred. No. 0.016;  
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTTGAAGGGCGCCGATTTGACATGCTTATGTCGAACCTGATGCTGAAG 597  
Db 298 HSNLSLGHVPRWPKNINKLQELDLSONFLAKEIGDAKFLHPLPSLLQLDLSFNFELQYVR 357

Qy 598 GGTTCGAGAGTCTCTTCCAAATCGACTGTCTACGACTGACGACGCTTTTGGAC 657  
Db 358 ASMNLSQARSSLSKSLKILIRGVFEKLSFNLSPLHNLQNEVLDTGNTFIKIANLSMF 417  
Qy 658 CAATGTTCCGATAGTGAAGCATGGCTTTCTTCCGAGAGATGTTCTTAAGCGAAAGTG 717  
Db 418 KQFKRLKVIDLSVKNISPSGDSEVFGCSNARTSVSYPQVLEQHLHYRYDKVANSR 477  
Qy 718 GCGGTATTGGGCTGACATTTCCGAGCTGTGTGTCGCAAGCAACTGCTTACGCTGG 777  
Db 478 KXKEASFMVSNNESCVYGGTLLDLSKNSIFEVKSSDPQHLFLKCLNLSNLSQTLNGSE 537  
Qy 778 GTAGACATGTTACAAATATATGAAAGCAAGTATCGTTGAGGCAAGCTTGGTCACAT 837  
Db 538 FQPLAEIRYDFSNRNLIDLIHSTAEEBLHKEVLIDISSNSHYFQSEGITMLNFTKMLKV 597  
Qy 838 GCTTTCAGGAGCCTCTAGTGTGTGCGCAATGGGGCGATGCGATTTCCCTGCT 897  
Db 598 LQKLMNNDNDISSSTRTMSESLRTLEFRGNHLDVLMREGDNRVYQLFKNILKLEBDI 657  
Qy 898 GCATTTGCTTTGTTTTCTTCTCGAGCGCTTACCGCCTGTCTTGATGAGCGCGTTCCA 957  
Db 658 SKNSLFLPSGVFDGMPNMLKNLSLAKNGLKSFWSKKLOCLKNLETLIDLSHNLTVPER 717  
Qy 958 AATCCCGGCACTGACACTTACTGCTGTACCAAGCGCTCCAAATCATGTGGAAGCC 1017  
Db 718 LSNCSLSLKNLILKNNQIRSLTKYFLQDAPQLRYLIDLSNNKIQMTQKTSFPEVNLNLM 777  
Qy 1018 GGCAGCTGCGCCAGAGCTGTCATCGGCTTACCAAGCTTGGGCTGCTTGAAG 1077  
Db 778 LLLHNRFLCTCAVFWFVWVWVHTVTIPLYLATDVTCVPAHKGQSVSLDYLTBELD 837  
Qy 1078 GACGGTTTTCATGAGAGATATGTGTGCTTGGCTTGGCTGCTATTACTCAGGCTTG 1137  
Db 838 TNLILFSLISVSLFLMVMWMTASHLYFMDVWYIYHFKAKIKGYORLISPDCYDAFIY 897  
Qy 1138 AATCAGACACATTAGTGCGCTCATGACTCCTGCGAAATTTGGCTGAACGTTTGGG 1197  
Db 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLBERDMLPGQVLENLSQSIQLSKTVFWM 957  
Qy 1198 AGGAGTCTCTCTTCTCAGGATAGAGAGATCTTCTGCGCACACAT 1245  
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILILEKPFQSKFLQLRK 1005

RESULT 9  
US-10-140-805-358  
; Sequence 358, Application US/10140805  
; Publication No. US20030207417A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C176  
; CURRENT APPLICATION NUMBER: US/10/140, 805  
; PRIOR FILING DATE: 2002-05-07

```

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRF
; ORGANISM: Homo Sapien
US-10-140-805-358

Query Match
  2.3%; Score 40.8; DB 12; Length 1049;
  Best Local Similarity 5.4%; Pred. No. 0.016;
  Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY ACCCTTGAAGGGGCGGCGCATTTGACATGAGTGTATGTCGAAACCTGATGCTGAAG 597
Db HNSNLSOHVPRPFWKINIKQELDLSONFLAKEIGDAKFLHPLSLQLDLSFPELQYVR 357
QY 598 GGTTCGACAGTTCCTTTCCAAACATCGATTCCTACGATCGACGACCGTTTTTGAAC 657
Db ASMNLSQAFSSLSKLILRIGVFEKLSFNLSPHLNQLNLEVLDTGNFIKIANLSMF 417
QY 658 CAATGTCGATAGTGAAGGATCGGCTTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
Db 418 KQFKRLKVIDLSYKISPSGDSSEVGFCSNARTSVSEYPOVLEOLHYFRYDKYASCRF 477
QY 718 GCGGTCATTTGGCGCATTTCCGACCTGCTGTCGCAACGACGCTTCATGCTGGG 777
Db 478 KKNKASFMVSNESCYRGOTLDLSKNSIFPVKSSDQHSFLKCLNLSNLSQTLNGSE 537
QY 778 GTAGACGATGTATCAATATATGAGCAAGTATGCTGTGGAGCAAGCTTTGGTCAAT 837
Db 538 FQPLALRLYLDPSNNRLDLHSTAFELHKLLEVLDISNSHYFQSGITMNLFTNKLKY 597
QY 838 GCTTTCAGGAGCGCTCTAGTGTCTGCGCCGAATGGGGCGCATTCCTCTCGCT 897
Db 598 LQKLMMNDNDISSSTRTWESLSRTLFRGNHLDVLMREGDNRYLQLFNKLKLEBLDI 657
QY 898 GCATTCGCTGTTTTTCTTCCTCGACGCTTACGCGCTGCTTCGATGAGCGCTTCCA 957
Db 658 SKNSLSFLPSGVGDMPNPKNLISLAKNGLSFSWKLQCLNLEVLDSHNDLTVPER 717
QY 958 AATCCCGGACAGTGCAGACTTCTGCTTACCAAGCGCTCCAAATACATGATGGAAGCC 1017
Db 718 LSNCSLSLKLILKNNQIRSLTYFLQDAFQLRYLDLSNNKIQIMQKTSPEVNLNKLK 777
QY 1018 GGGCAGCTGCCACGAGCTGTTCCATCGGCTTACACGCTTGGCGCTGCTTGAAG 1077
Db 778 LLHNRFLCTCDVAVFWVWVNNHTEVTIPLYLATDVTCVGGAHKGGSVISLDYTCBLD 837
QY 1078 GACGCTTTTATAGAGGAGATATGTTGTTGGCTTCGCTGCTATTCACGAGCTTG 1137
Db 838 TNLILPSLSISVSLFLWMMNTASHLYFWDVWYIYHFCAKIKGYORLISPDCCYDAFIV 897
QY 1138 AATCCAGACACATTAAGTGGGCTCATGATCTCTCGCAAAATTTGGTGAACGTTTGGG 1197
Db 898 DTDPAVTEVNLVLAELVAKLEDPREKHNLCLERBDWLPQGVLENLSQSILSKTYFVM 957
QY 1198 AGGAGTCTTCTCTTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
Db 958 TDYKAKTENFKIAFYLSHQRLMDKVDVILLIFLEKPFQSKFLQLRK 1005

RESULT 10
US-10-140-864-358
; Sequence 358, Application US/10140864
; Publication No. US20030207419A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flivaro, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geertsema, Mary E.
```

```

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACTS ENCODING THE SAME
; FILE REFERENCE: P330R1C184
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRF
; ORGANISM: Homo Sapien
US-10-140-864-358
```

```

Query Match
  2.3%; Score 40.8; DB 12; Length 1049;
  Best Local Similarity 5.4%; Pred. No. 0.016;
  Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;
```

```

QY 538 ACCCTTGAAGGGGCGGCGCATTTGACATGAGTGTATGTCGAAACCTGATGCTGAAG 597
Db 298 HNSNLSOHVPRPFWKINIKQELDLSONFLAKEIGDAKFLHPLSLQLDLSFPELQYVR 357
QY 598 GGTTCGACAGTTCCTTTCCAAACATCGATTCCTACGATCGACGACCGTTTTTGAAC 657
Db 358 ASMNLSQAFSSLSKLILRIGVFEKLSFNLSPHLNQLNLEVLDTGNFIKIANLSMF 417
QY 658 CAATGTCGATAGTGAAGGATCGGCTTTCCGAGAGATGTTCTTAAGCCGAAGTG 717
Db 418 KQFKRLKVIDLSYKISPSGDSSEVGFCSNARTSVSEYPOVLEOLHYFRYDKYASCRF 477
QY 718 GCGGTCATTTGGCGCATTTCCGACCTGCTGTCGCAACGACGCTTCATGCTGGG 777
Db 478 KKNKASFMVSNESCYRGOTLDLSKNSIFPVKSSDQHSFLKCLNLSNLSQTLNGSE 537
QY 778 GTAGACGATGTATCAATATATGAGCAAGTATGCTGTGGAGCAAGCTTTGGTCAAT 837
Db 538 FQPLALRLYLDPSNNRLDLHSTAFELHKLLEVLDISNSHYFQSGITMNLFTNKLKY 597
QY 838 GCTTTCAGGAGCGCTCTAGTGTCTGCGCCGAATGGGGCGCATTCCTCTCGCT 897
Db 598 LQKLMMNDNDISSSTRTWESLSRTLFRGNHLDVLMREGDNRYLQLFNKLKLEBLDI 657
QY 898 GCATTCGCTGTTTTTCTTCCTCGACGCTTACGCGCTGCTTCGATGAGCGCTTCCA 957
Db 658 SKNSLSFLPSGVGDMPNPKNLISLAKNGLSFSWKLQCLNLEVLDSHNDLTVPER 717
QY 958 AATCCCGGACAGTGCAGACTTCTGCTTACCAAGCGCTCCAAATACATGATGGAAGCC 1017
Db 718 LSNCSLSLKLILKNNQIRSLTYFLQDAFQLRYLDLSNNKIQIMQKTSPEVNLNKLK 777
QY 1018 GGGCAGCTGCCACGAGCTGTTCCATCGGCTTACACGCTTGGCGCTGCTTGAAG 1077
Db 778 LLHNRFLCTCDVAVFWVWVNNHTEVTIPLYLATDVTCVGGAHKGGSVISLDYTCBLD 837
QY 1078 GACGCTTTTATAGAGGAGATATGTTGTTGGCTTCGCTGCTATTCACGAGCTTG 1137
Db 838 TNLILPSLSISVSLFLWMMNTASHLYFWDVWYIYHFCAKIKGYORLISPDCCYDAFIV 897
QY 1138 AATCCAGACACATTAAGTGGGCTCATGATCTCTCGCAAAATTTGGTGAACGTTTGGG 1197
Db 898 DTDPAVTEVNLVLAELVAKLEDPREKHNLCLERBDWLPQGVLENLSQSILSKTYFVM 957
QY 1198 AGGAGTCTTCTCTTCAGGAGATAGAGAGATCTTCTGGGACACAT 1245
Db 958 TDYKAKTENFKIAFYLSHQRLMDKVDVILLIFLEKPFQSKFLQLRK 1005
```

Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILLIFLEKPFQSKFLQLRK 1005

## RESULT 11

US-10-140-923-358

Sequence 358, Application US/10140923  
Publication No. US20030207355A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C188  
CURRENT APPLICATION NUMBER: US/10/140,923  
CURRENT FILING DATE: 2002-05-07  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-140-923-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;

Best Local Similarity 5.4%; Pred. No. 0.016; Mismatches 473; Indels 0; Gaps 0;

Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTGAAGGGGCGCGCATTTGACATGATGCTGAACCTGATGCTGAAG 597  
Db 298 HNSLSQHVPPRPMFNKINQLQELDSQFLAKEIGDAKFLHFLPSLIQLDLSFNFELQYR 357  
Qy 598 GGTTCGGCAGGTTCTTCCAAACATGACCTGCTTACAGCTGACAGCCGTTTGGAC 657  
Db 358 ASNMLSQAFSSLSKSLKILIRGVFKELKSFNLSPLNLQNLVLDLGTNFIKIANLSMF 417  
Qy 658 CAATGTTCCGATGATGACGAGATGCGCTTTTCCGAGAGATGTTCTTAAGCCGAAAGTG 717  
Db 418 KQFRLKLVLDLSVKNLSPSGDSSEVFCNSARTSVSEYPOVLEQLHYFYDYKARSQRF 477  
Qy 718 GCGGTCATGGCGCTGCGATTCGCGACTGCTGGTGCAACGAACTGCTTACGCTGAGG 777  
Db 478 KKEASMSVNSCYVGGTLDLSKNSIFVYKSSDPQHLFLKLNLSGMLISQTLNGSE 537  
Qy 778 GTAGACATGTTACATATATGAGCAAGTATGCTGTTGAGGCAAGCTTGGTGCACAT 837  
Db 538 FQPLAELRYLDLFSNNRDLHSTAFBELHLEVLDISNSHYPGSEGITMLNFTKNLKV 597  
Qy 838 GCTTTCGAGGACGCTCTAGTGTGTCGCGGAAATGGGGCGATGATTTCTCTGCTGCT 897  
Db 598 LQKLMDNDNDISSSTRTSESLRLTEFRGNHLDVLMRGDNRYLQLFKNLKLLELDI 657  
Qy 898 GCATTCGCTGTTGTTTCTTCTCGAGCTTACGCGCTGCTTCTGATGAGAGCCGCTTCCA 957  
Db 658 SKNSLSLBPBGVFBGMPNKLKNSLAKNGLSKWSKLQCLKLETLIDLSHNOITLTPER 717  
Qy 958 AATCCCGGACAGTCACTTACTGCTTACCAAGCGCTCCAAATCATGATGGAAGCC 1017  
Db 718 LSNCSRLKLNILKNNQIRSLTKVFLQDAPQLRLYLDLSNKNQIQKTSFPENVLNKLKM 777

Qy 1018 GCGAGCTGCCACGAGCTTCTTCATGCGCTTACACAGCTTGGCGCTTCTGAAG 1077  
Db 778 LLLHNRFLCTCAVAFVWVWVHTVITPLATDVTCVGBAHKGSVSLDYTELDEL 837  
Qy 1078 GACGTTTTCATGAGCAGATATGTTGGCTTCCGCTGCTGCTTACTTACAGCCCTTG 1137  
Db 838 TNLIFSLISVSLFLVMMWMTASHLYFWDVWVYIYHCKAKIKGYRLISPDCCVDAPFVY 897  
Qy 1138 AATCAGACACATTAGTGGCTCATGACTCTCGCAAAATTGGCTGAACGCTTTCGGG 1197  
Db 898 DTPDPVTEVWVLAELVAKLEDPREKHNLCLEERDMLPGQPVLENLSQSISQTSKTVFVW 957  
Qy 1198 AGGAGCTCTTCTTCCGAGGATGAGAGATCTTCTGCGACACAT 1245  
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILLIFLEKPFQSKFLQLRK 1005

## RESULT 12

US-10-141-756-358

Sequence 358, Application US/10141756  
Publication No. US20030207355A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C200  
CURRENT APPLICATION NUMBER: US/10/141,756  
CURRENT FILING DATE: 2002-05-08  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-141-756-358

Query Match 2.3%; Score 40.8; DB 12; Length 1049;

Best Local Similarity 5.4%; Pred. No. 0.016; Mismatches 473; Indels 0; Gaps 0;

Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTGAAGGGGCGCGCATTTGACATGATGCTGAACCTGATGCTGAAG 597  
Db 298 HNSLSQHVPPRPMFNKINQLQELDSQFLAKEIGDAKFLHFLPSLIQLDLSFNFELQYR 357  
Qy 598 GGTTCGGCAGGTTCTTCCAAACATGACCTGCTTACAGCTGACAGCCGTTTGGAC 657  
Db 358 ASNMLSQAFSSLSKSLKILIRGVFKELKSFNLSPLNLQNLVLDLGTNFIKIANLSMF 417  
Qy 658 CAATGTTCCGATGATGACGAGATGCGCTTTTCCGAGAGATGTTCTTAAGCCGAAAGTG 717  
Db 418 KQFRLKLVLDLSVKNLSPSGDSSEVFCNSARTSVSEYPOVLEQLHYRYDYKARSQRF 477  
Qy 718 GCGGTCATGGCGCTGCGATTCGCGACTGCTGGTGCAACGAACTGCTTACGCTGAGG 777  
Db 478 KKEASMSVNSCYVGGTLDLSKNSIFVYKSSDPQHLFLKLNLSGMLISQTLNGSE 537

QY	778	GTAGACAGATGTTACAATATATAAGGAACAAGTATCGTGTGGAGGCAAGCTTTGGCACAT	83.7
Db	538	FORLAELRYLDLFSNNRDLHLSTAFEBELHKLEVDLISNSHYIFOSEGITHHLANTTKILXY	59.7
QY	838	GCTTTACAGGACGCGCTCTAGTGTGCGGCCGAATGGGGGCGATGCGATTTCTCTGCT	89.7
Db	598	LQKLMNDNDNDISSSTRMSESLRTELEFRGNHLDVLRBEGDNRVLQLEFKMLKLEELDI	65.7
QY	898	GCATTCTGCTGTTTTCTTCTCTCGAGCGTTAAGCGCTGCTTGATGAGGCGCGTCCCA	95.7
Db	658	SKNSLSLPSGVFEGMPRLKNLSLAKGLSFSWKILQCLKNLETLDSHNLTLTPER	71.7
QY	958	AATCCCGGACAGTGCAGCACTTACTTGCTTACCAAGGCGCTCAATACATGTGAAAACC	101.7
Db	718	LSNCSRSRLKQNILKQNGRLSLTKYFLDQAFQRLYLDLSNKKIQMLQKTSFPEVNLNLMK	77.7
QY	1018	GGGCACTGCCACCGGAAGCTTTCCATCGCGCTTTAACAGGTTGGCGCGCTTTGAAG	107.7
Db	778	LLLHNHNFLLCTCDADVFWFVWVNNHNEVLTIPYATLDTTGVGRANKGQSVLSLDLYCELDL	83.7
QY	1078	GACGGTTTTCATGAGCGAGATATTGTGTGGCTTCGCGCTGCGCTATTAACAGGCTTG	113.7
Db	838	TNLLIFSLSISVSLEFWMVMTASHLYFMDVWYIHFCSAKIKGYRLSPDCCDAFLVY	89.7
QY	1138	AAATCAGACACATTAGGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACCGTTTCGGG	119.7
Db	898	DTKQPAVTEWVLAELVAKLEDPREKHFIKLEBRDMLPGQFVLELISQISLDSKTVFVM	95.7
QY	1198	AGGAGTCTCTCTCTCAGGATGAGAGCATCTTTCGGGCACACAT	124.5
Db	958	TDKAKTENFKIAFLYLSHQRLMDKENVDTYIIIFLEKRPQSKFLQRL	100.5

```

RESULT 13
US-10-141-759-358
; Sequence 358, Application US/10/141759
; Publication No. US20030207361A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C197
; CURRENT APPLICATION NUMBER: US/10/141,759
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-141-759-358

```

Query Match 2.3%; Score 40.8; DB 12; Length 1049;  
 Best Local Similarity 5.4%; Pred. No. 0.016;  
 Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0.

Oy 538 ACCCTTGAAAGGGGCGCTCCATTTCACATGGTAGCTTATGGTCGAAACCTGATGCTGAAG 597

Dh	298	HSNSLGHVPRMFWKINIKLQELDLSQNFLLKIGDAKFHFLPRLSLDLSFNFELQVLR	357
QY	598	GGTTCCGACAGTTCCTTCCACAACTGACTTGTCTACAGCTGCAGACCGTTTGTGAC	657
Dh	358	ASNNLSQAFSSLSKLKLRIRGVVFEXELKSFNLSPLNLONLEVLDTGTFIKIANLSMF	417
QY	658	CAATGTTCCGATGTGGACGAGATCCGTTCTTCCGAGGATGTCTTAAGCCGAAGTG	717
Dh	418	KQFKRLKVLDSLNNKISPSGDSSEVGFCSNARISVESYERQVLEQLHYHFRYDKABRCRF	477
QY	718	GCGGTCATTCGCGCTGGCATTTCCGACCTCGTGATGGCAACGACTTCATGCTGGG	777
Dh	478	KKNBASPMVSNESCYKKYGGTLDLSKNSIFPVKSSDFHLSFLKYNLSGNLISQTLNGSE	537
QY	778	GTGACGACTTTCACATATATATGAAGCAAGTACGTCTGTGGAGCAAGCTTTGTCACAT	837
Dh	538	FOPLAELRYLDFSNRLDLHSHAFEBLHLEVLDISNSHYPOSSEGITMHLNFTKLV	597
QY	838	GCTTTCAGGAGCGTCTAGTGTGTCGGCCGAATGAGGGCGATGCGATTTCCTCGTCT	897
Dh	598	LQKLMMDNDISSSTRTMESESLRTLEFFGNHLDVIMRGDNRYLDFPKLKLLELDI	657
QY	898	GCATTCGCTGTTTCTTCTCTCGACGCTTAAGCGGCTGTCTTCGATGAGCGCTTCCA	957
Dh	658	SKNSLSLPSGVDPDMRPMLKINLSANGKLSWKLCQLKULETLDLSHQTLTVBER	717
QY	958	AATCCGGCAGCTGCAGACTTAATGTTGGTTACCAAGCGCTCCATATCATGTGAAAGCC	1017
Dh	718	LSNCSRSLLKMLIKMNOIRSLTKYFLODAFOLRYLDSLNNKIQMIOKTSRPENVLNNLKM	777
QY	1018	GGGAGCTGCACCGAAGCTGTTCCACGGGTTTACACGGTGGCGCTGTTCTAAG	1077
Dh	778	LILHNNRFLTCDAAVFWFVMMNHEVITPYLADVTGVRGAKGQSVISLDLYTCELDD	837
QY	1078	GACGGTTTCATGACGACAGATATTTGTTGGCTTCGCTGCGCTATTAATCAGACCTTG	1137
Dh	838	TNLLIFSLISVSILFWMVMTASHLYWVDWVIYHPEKAKIKGVORLISPDCVDAPILV	897
QY	1138	AAATCAGGACACATTAGGTGGGCTCATGACTCCTGGCAAAATTTGGCTGAACCGTTTGGG	1197
Dh	898	DTKDPATVEWVLAEVLAKLEDPREKHNNLCEERDMLPGRVLENLSQISLSSKTVFVM	957
QY	1198	AGGAGACTTCTCTTCGAGGATGACAGAGATCTTTCGGCAGACACAT	1245
Dh	958	TDKAKENFKIAFYLSHORLMBKDVVILIIFLEKFPQSKFLQDLRK	1005

RESULT 14  
US-10-123-155-358  
Sequence 358, Application US/10123155  
Publication No. US20030068794A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P9330RIC30



Oy 1198 AGGAGTCTTCTTCAGGATAGAGGATCTTCTGGGACACAT 1245  
Db 958 TDKYAKTENFKIAFYLSHQRLMDKVDVILIFLEKPFQSKFQLQLRK 1005

Search completed: November 23, 2003, 21:22:48  
Job time : 560.252 secs



PT untranslatable nucleic acid homologous to a gall disease-causing gene  
XX  
PS Claim 2, Page 45; 49pp; English.

XX The patent discloses a method for producing plants resistant to gall  
CC disease, induced by Agrobacterium infection, by transformation with a  
CC bacterial resistance (BR) construct capable of eliciting co-suppression.  
CC The construct encodes an untranslatable RNA molecule, that is highly  
CC homologous to the tumor or gall disease-causing genes *iaam*, *iaah* and  
CC *ipt* of *Agrobacterium tumefaciens*. These oncogenes are modified by  
CC introduction of premature termination codons or frameshift mutations, to  
CC inhibit their expression. This method is used to reduce susceptibility  
CC of plants to gall diseases. The present DNA sequence is the *Agrobacterium*  
CC tumefaciens T-DNA (transferred DNA) gene *iaam*, required for auxin  
CC production. It encodes an enzyme that converts tryptophan into indole  
CC acetamide, essential for auxin synthesis. The overproduction of these  
CC plant growth hormones results in gall formation.

XX Sequence 2268 BP; 543 A; 520 C; 606 G; 599 T; 0 other;

Query Match 100.0%; Score 2268; DB 21; Length 2268;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTCACCTTCACTCTCTGTAACCACTGCGATCATCTCCCAACCAAAATGCTGAT 60
DB 1 ATGTCACCTTCACTCTCTGTAACCACTGCGATCATCTCCCAACCAAAATGCTGAT 60
QY 61 CTGACATAGTGTGATGAGGAGATGGAACCGAGGGTTTCCAGTCCCTTTAGAA 120
DB 61 CTGACATAGTGTGATGAGGAGATGGAACCGAGGGTTTCCAGTCCCTTTAGAA 120
QY 121 CGAAGACCTTCAAGGGAGAGGATTAATCAAGCTTCAACCGAGTCCAGCGCTGGTTA 180
DB 121 CGAAGACCTTCAAGGGAGAGGATTAATCAAGCTTCAACCGAGTCCAGCGCTGGTTA 180
QY 181 GCTTGAAGAGGCGGCGGATGTCGCTCCCGAGATCTCAGTGTGGAAGGTGCA 240
DB 181 GCTTGAAGAGGCGGCGGATGTCGCTCCCGAGATCTCAGTGTGGAAGGTGCA 240
QY 241 GTTCTCTCCGCTTATATCTATATGCAAGAAATTTGCGGCGGATTAATCGAAA 300
DB 241 GTTCTCTCCGCTTATATCTATATGCAAGAAATTTGCGGCGGATTAATCGAAA 300
QY 301 CTTTGGCGCGGCGACAGTGAAGTGTCTCGTTCATCACTTGAACCAATTTGCATG 360
DB 301 CTTTGGCGCGGCGACAGTGAAGTGTCTCGTTCATCACTTGAACCAATTTGCATG 360
QY 361 GATTTTCCGAGACAACTAATTCAGAGCCCTGTTTTGCTGAGCGGTAAGATGCA 420
DB 361 GATTTTCCGAGACAACTAATTCAGAGCCCTGTTTTGCTGAGCGGTAAGATGCA 420
QY 421 CCGATTGATCTTAGTCAATTCGTCGTCATTTCAATCTTAAGATGCGGCTTTGCAAC 480
DB 421 CCGATTGATCTTAGTCAATTCGTCGTCATTTCAATCTTAAGATGCGGCTTTGCAAC 480
QY 481 CTGCGCAATGCGCTGTAGCAAGATGCGAGATGAATGCGTTACCGGGTTTACCAATAC 540
DB 481 CTGCGCAATGCGCTGTAGCAAGATGCGAGATGAATGCGTTACCGGGTTTACCAATAC 540
QY 541 CTTGAAGAGGGCCGTCGATTTGACATGTAGCTTAATGTTGGAACCTGATGCTGAAGGT 600
DB 541 CTTGAAGAGGGCCGTCGATTTGACATGTAGCTTAATGTTGGAACCTGATGCTGAAGGT 600
QY 601 TCGGAGAGTCTCTTCCAAATGCACTTCTCTACAGCTGACAGCCGTTTTTTGACCA 660
DB 601 TCGGAGAGTCTCTTCCAAATGCACTTCTCTACAGCTGACAGCCGTTTTTTGACCA 660
QY 661 TGTTCGATAGTGAACGATCGGCTTTTCCGAGAGATTTCTTAAGCCGAAAGTGCG 720
DB 661 TGTTCGATAGTGAACGATCGGCTTTTCCGAGAGATTTCTTAAGCCGAAAGTGCG 720
```

```
QY 721 GTCATTGGCGCTGGCATTTCCGAGCTGTCGTGGCAACCAACGCTTCAATGCTGGGTA 780
DB 721 GTCATTGGCGCTGGCATTTCCGAGCTGTCGTGGCAACCAACGCTTCAATGCTGGGTA 780
QY 781 GACGATGTTAACAATATATGAGCAAGATGATCGTGTGAGGCAAGCTTTGTCATGCT 840
DB 781 GACGATGTTAACAATATATGAGCAAGATGATCGTGTGAGGCAAGCTTTGTCATGCT 840
QY 841 TTGAGGAGCGCTCTTGTGTGTGTGCGGCAAAATGGGGCCGATGCGATTTCTCTGTGCA 900
DB 841 TTGAGGAGCGCTCTTGTGTGTGTGCGGCAAAATGGGGCCGATGCGATTTCTCTGTGCA 900
QY 901 TTCTGCTGTTTTTCTCTCTCGAGCGTTACGCGCTGTCTCATGTAGAGCGCTTCCCAAT 960
DB 901 TTCTGCTGTTTTTCTCTCTCGAGCGTTACGCGCTGTCTCATGTAGAGCGCTTCCCAAT 960
QY 961 CCGGCAACAGTCGACACTTAATTGTTCTACCAAGGCGTCCAAATCATGTGAAAGCCGG 1020
DB 961 CCGGCAACAGTCGACACTTAATTGTTCTACCAAGGCGTCCAAATCATGTGAAAGCCGG 1020
QY 1021 CAGCTGCCACCGAAGCTGTTCCATCCGCTTTACACAGGTTGGCGTCTTTGAAAGAC 1080
DB 1021 CAGCTGCCACCGAAGCTGTTCCATCCGCTTTACACAGGTTGGCGTCTTTGAAAGAC 1080
QY 1081 GGTTCATGAGCGAGATATGTTGTTGGCTTGCCTGTGCTATTAATCTCAGGCTTGAAA 1140
DB 1081 GGTTCATGAGCGAGATATGTTGTTGGCTTGCCTGTGCTATTAATCTCAGGCTTGAAA 1140
QY 1141 TCAAGACATTAAGTGGGCTCATGATCTCTGGCAAAATTTGCTGAAACCGTTTCGGAG 1200
DB 1141 TCAAGACATTAAGTGGGCTCATGATCTCTGGCAAAATTTGCTGAAACCGTTTCGGAG 1200
QY 1201 GAGTCTCTCTCTCAGGATATGAGAGATCTTTCTGGGCAACATCTCTGCTGTGTTAA 1260
DB 1201 GAGTCTCTCTCTCAGGATATGAGAGATCTTTCTGGGCAACATCTCTGCTGTGTTAA 1260
QY 1261 CATGAGATTTTCTCATGATGTTGGAACCTTAATGAGATATGGAATGAGATCTGGCGGG 1320
DB 1261 CATGAGATTTTCTCATGATGTTGGAACCTTAATGAGATATGGAATGAGATCTGGCGGG 1320
QY 1321 TTTGTCGATTTTGAAGCGGGTTTATGAGATCTCTCGCTTGTGATCAACGAGATAT 1380
DB 1321 TTTGTCGATTTTGAAGCGGGTTTATGAGATCTCTCGCTTGTGATCAACGAGATAT 1380
QY 1381 GAAGAAATCAGGAGATGTCCTGGAAGAAATTCAGAACTTCCAGTCCGATCGCATCT 1440
DB 1381 GAAGAAATCAGGAGATGTCCTGGAAGAAATTCAGAACTTCCAGTCCGATCGCATCT 1440
QY 1441 GAAGTGTAAAGGATGTCGTGAGCAGCGCATATGCCATGTTCAAGTCAGGCGATT 1500
DB 1441 GAAGTGTAAAGGATGTCGTGAGCAGCGCATATGCCATGTTCAAGTCAGGCGATT 1500
QY 1501 CAGAGAAAAAGCAAAAAATTAAGATTAAGCTTAAAGCGGATATCTGAACCTTTATGAT 1560
DB 1501 CAGAGAAAAAGCAAAAAATTAAGATTAAGCTTAAAGCGGATATCTGAACCTTTATGAT 1560
QY 1561 AAGTGTGTGTCATCTGATCTGCAATATCAACTCAGGATTTGCTGATGATCGAT 1620
DB 1561 AAGTGTGTGTCATCTGATCTGCAATATCAACTCAGGATTTGCTGATGATCGAT 1620
QY 1621 ACCAATATTTTTCAGGACACAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATCG 1680
DB 1621 ACCAATATTTTTCAGGACACAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATCG 1680
QY 1681 TCAAAACTTCTCTGATGATGACGAAAAATTTGTTGATGACCATATCTCCGCTTTGT 1740
DB 1681 TCAAAACTTCTCTGATGATGACGAAAAATTTGTTGATGACCATATCTCCGCTTTGT 1740
QY 1741 GTTCTTATGAGCGGATTCGAAAAAGCATGTATGCTGTCATATGATGTCAGATCCG 1800
DB 1741 GTTCTTATGAGCGGATTCGAAAAAGCATGTATGCTGTCATATGATGTCAGATCCG 1800
QY 1801 AATGTAAGGTTATAGTCTCACTTAATTAATGAGGAGACGATCCCAAGAGCTTTG 1860
```

```

Db 1801 AATGTAAGAGTCTAGTCTCATCAGTTATACATGGAGACGACTCCCAAGCTGTG 1860
Qy 1861 GCGGTCCCCGCAAAAAAGGCGATTATGCTCTGGGGAGCCCAATTTGAGATCTTTC 1920
Db 1861 GCGGTCCCCGCAAAAAAGGCGATTATGCTCTGGGGAGCCCAATTTGAGATCTTTC 1920
Qy 1921 CCGGCGTTTGCAGCAGCCTATTTCTGCTGCGCTGATTAAGCAACCAAAATTTATTC 1980
Db 1921 CCGGCGTTTGCAGCAGCCTATTTCTGCTGCGCTGATTAAGCAACCAAAATTTATTC 1980
Qy 1981 CATGATTTGCTTACAGAGCAATGCGGGGGAGCTTTCAACTCAACCGGCGTGTGAG 2040
Db 1981 CATGATTTGCTTACAGAGCAATGCGGGGGAGCTTTCAACTCAACCGGCGTGTGAG 2040
Qy 2041 GATTTTATTTCTAAGAACTTTCTTTCAGACACTGACACGCGTATATGATACCGAGTT 2100
Db 2041 GATTTTATTTCTAAGAACTTTCTTTCAGACACTGACACGCGTATATGATACCGAGTT 2100
Qy 2101 TACTTGGCGGTTGCAAGTTGTTCTTTCACAGTGATGGGTGAGGCTGATTTACAGACC 2160
Db 2101 TACTTGGCGGTTGCAAGTTGTTCTTTCACAGTGATGGGTGAGGCTGATTTACAGACC 2160
Qy 2161 GCGGTGTAACGCGCTGCTGCAATTTATCCAAATTTGTGAGGCAATTTTGCAAAAGGCAAT 2220
Db 2161 GCGGTGTAACGCGCTGCTGCAATTTATCCAAATTTGTGAGGCAATTTTGCAAAAGGCAAT 2220
Qy 2221 CCTCTGAAACACTCTTGGAAAGATATATACCTCCGCGTGAATTTG 2268
Db 2221 CCTCTGAAACACTCTTGGAAAGATATATACCTCCGCGTGAATTTG 2268

```

## RESULT 2

ID AAN50182 standard; DNA; 24596 BP.

AAN50182;

XX 25-MAR-2003 (updated)

DT 17-OCT-1991 (first entry)

XX Complete nucleotide sequence of the T-DNA region of the octopine Ti

XX plasmid pTi15955.

XX Plant vector; transformation-inducing principle (TiP) gene;

XX octopine Ti plasmid; ss.

XX Agrobacterium tumefaciens ATCC 15955.

XX EPI45338-A.

XX 19-JUN-1985.

XX 16-NOV-1984; 84BP-0307969.

XX 18-NOV-1983; 83US-0553786.

XX (AGRK) AGRIGENETICS RES ASSOC LTD.

XX (LUBR) LUBRIZOL GENETICS INC.

XX Barker RF, Kemp JD;

XX WPI, 1985-148223/25.

XX New DNA vectors contg. T-DNA sequence of octopine Ti plasmid -

XX for expression in plant cells to confer desirable properties to

XX plants and their cells

XX Claim 28; Fig 1; 87bp; English.

XX The inventors claim a vector contg. a transformation-inducing

XX principle (TiP) gene from Ti plasmid pTi15955. The sequence of the

XX T-DNA of the octopine-type Ti plasmid has fourteen open reading

CC frames bounded by eukaryotic promoters, ribosome binding sites, and  
 CC polyadenylation sites. With the vectors, expression of structural  
 CC foreign genes in plant cells is promoted. The gene esp. encodes an  
 CC insecticidal toxin identical to or derived from the crystal protein  
 CC of *Bacillus thuringiensis*.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX

SQ Sequence 24596 BP; 6534 A; 5510 C; 5793 G; 6759 T; 0 other;  
 Query Match 98.4%; Score 2230.6; DB 6; Length 24596;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```

Qy 1 ATGTGACCTTCACTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGTGTGAT 60
Db 5809 ATGTGACCTTCACTCTCTGATTAACAGTGGATCATCTCCCAACCAAAATGTGTGAT 5868
Qy 61 CTGACAAATGTCGATTAAGGCGGATGAATGACCCGAGGGTTTCGATGCTTTTAGAA 120
Db 5869 CTGACAAATGTCGATTAAGGCGGATGAATGACCCGAGGGTTTCGATGCTTTTAGAA 5928
Qy 121 CGAAGAGCTTTCTAAGGGAAGAGATTAATCAAGCTCCACCGAGTGCAGCGCTGGGTTA 180
Db 5929 CGAAGAGCTTTCTAAGGGAAGAGATTAATCAAGCTCCACCGAGTGCAGCGCTGGGTTA 5988
Qy 181 GCTTGCAAAAGGCTGGCGATGATGCTTCCCGAGATCTCAGCTGTTGAAAAGTATGA 240
Db 5989 GCTTGCAAAAGGCTGGCGATGATGCTTCCCGAGATCTCAGCTGTTGAAAAGTATGA 6048
Qy 241 GTTCTCTCGCTTATATCTATATTGCAAAAGAAATCTGGGGCGGATTAATTTAGAA 300
Db 6049 GTTCTCTCGCTTATATCTATATTGCAAAAGAAATCTGGGGCGGATTAATTTAGAA 6108
Qy 301 CCTTGGCGCGGCGCAACAGTGAATGCTCTGCTTCCATGATGACTTGAACACATTTGCAATG 360
Db 6109 CCTTGGCGCGGCGCAACAGTGAATGCTCTGCTTCCATGATGACTTGAACACATTTGCAATG 6168
Qy 361 GATTTCTCGAAGCAACATTAATCCAGGCGCTTTTCTGAGCGGTTAAAGATGTGA 420
Db 6169 GATTTCTCGAAGCAACATTAATCCAGGCGCTTTTCTGAGCGGTTAAAGATGTGA 6228
Qy 421 CCGATTAATCTAGTCAATTTGCGGCAATTTCAATCTTAAGACTGCGGCTTTTGAAC 480
Db 6229 CCGATTAATCTAGTCAATTTGCGGCAATTTCAATCTTAAGACTGCGGCTTTTGAAC 6288
Qy 481 CTGCAATGCGGCTGTGACGAAGATGCGACGATGAATGCGTTACCGGTTTACATTAAC 540
Db 6289 CTGCAATGCGGCTGTGACGAAGATGCGACGATGAATGCGTTACCGGTTTACATTAAC 6348
Qy 541 CTTGAAGGGGCGGTCGATTTGACATGATGCTTAATGTCGAACCTGATGCTGAAGGGT 600
Db 6349 CTTGAAGGGGCGGTCGATTTGACATGATGCTTAATGTCGAACCTGATGCTGAAGGGT 6408
Qy 601 TCGGACAGTTCTTTTCCAAACAAATCGACTTCTTAAGCTGACGACCGTTTGTGACCA 660
Db 6409 TCGGACAGTTCTTTTCCAAACAAATCGACTTCTTAAGCTGACGACCGTTTGTGACCA 6468
Qy 661 TGTTCGATAGTGAACGATTCGGCTTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 720
Db 6469 TGTTCGATAGTGAACGATTCGGCTTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 6528
Qy 721 GTTCATTTGGGCTGCGATTTCCGACCTCGTGTGCAAAAGAACTGCTTATGCTGGGGTTA 780
Db 6529 GTTCATTTGGGCTGCGATTTCCGACCTCGTGTGCAAAAGAACTGCTTATGCTGGGGTTA 6588
Qy 781 GACGATGTACATATATGAACGAAGTATCGTGTGAGGCAAGCTTTGTGACATGCT 840
Db 6589 GACGATGTACATATATGAACGAAGTATCGTGTGAGGCAAGCTTTGTGACATGCT 6648
Qy 841 TTCAGGACGCTCTTATGATGCTGTCGCAAAATGCGGCGATGCGATTTCTCTGCTGCA 900
Db 6649 TTCAGGACGCTCTTATGATGCTGTCGCAAAATGCGGCGATGCGATTTCTCTGCTGCA 6708

```

QY 901 TTGCGCTGTTTTCTTCTCGAGCGTTAGGCGCTGTCCTGATGAGCGCGTTCCCAAT 960  
 DB 6709 TTTGCTGTTTTCTTCTCTCGAGCGTTAGGCGCTGTCCTGATGAGCGCGTTCCCAAT 6768  
 QY 961 CCGGCGCAGTCCGACCTTACTTGTGTACCAAGCGCTCCCAATCATGTGAAAAGCCGGG 1020  
 DB 6769 CCGGCGCAGTCCGACCTTACTTGTGTACCAAGCGCTCCCAATCATGTGAAAAGCCGGG 6828  
 QY 1021 CAGTCCGACCGAAGCTTTCATCGCGTTTACAAGCTTGGCGCTGCTTCTGAAAGAC 1080  
 DB 6829 CAGTCCGACCGAAGCTTTCATCGCGTTTACAAGCTTGGCGCTGCTTCTGAAAGAC 6888  
 QY 1081 GGTTCATAGAGAGATTTGTTGGCTTCGCTGCTGCTTCTGAGCGCTTGA 1140  
 DB 6889 GGTTCATAGAGAGATTTGTTGGCTTCGCTGCTGCTTCTGAGCGCTTGA 6948  
 QY 1141 TCAGACACATTAAGTGGGCTCATGACTCGTGGCAATTTGGCTGAACCGTTTCGGAGG 1200  
 DB 6949 TCAGACACATTAAGTGGGCTCATGACTCGTGGCAATTTGGCTGAACCGTTTCGGAGG 7008  
 QY 1201 GAGTCTTCTCTCGAGGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGTGT-AA 1259  
 DB 7009 GAGTCTTCTCTCGAGGATAGAGAGATCTTTCTGGGCAACATCTCTGCTGTGTAA 7068  
 QY 1260 ACATGAGTTTCTCTCATGATTGGGACCTTATTCAGCTAATGGGAATAGGATCTGGCGG 1319  
 DB 7069 ACATGAGTTTCTCTCATGATTGGGACCTTATTCAGCTAATGGGAATAGGATCTGGCGG 7127  
 QY 1320 GTTTGCTCAGTTTGTGAAAGCGGTTTATTTAGATCTCTCGCTGCTCATCAACGAGTA 1379  
 DB 7128 GTTTGCTCAGTTTGTGAAAGCGGTTTATTTAGATCTCTCGCTGCTCATCAACGAGTA 7187  
 QY 1380 TGAAGAAATCAGCGGATGTGCTTGAAGATTTCAAACTTCAAGTGGATCGATC 1439  
 DB 7188 TGAAGAAATCAGCGGATGTGCTTGAAGATTTCAAACTTCAAGTGGATCGATC 7247  
 QY 1440 TGAAGTGTTAAGCGTGTCTGTGAGCCAGCGCATTTGCAATTTCAAGTCAAGGCGAT 1499  
 DB 7248 TGAAGTGTTAAGCGTGTCTGTGAGCCAGCGCATTTGCAATTTCAAGTCAAGGCGAT 7307  
 QY 1500 TCAGAGGAAAGACAAATAAAGATTAAGCTTAAAGCGGATATCTGAACTTTATGA 1559  
 DB 7308 TCAGAGGAAAGACAAATAAAGATTAAGCTTAAAGCGGATATCTGAACTTTATGA 7367  
 QY 1560 TAAAGTGTGTGATCATCTGAGCTCGCAATATCAACTGAGGATTTGCTGATCGCA 1619  
 DB 7368 TAAAGTGTGTGATCATCTGAGCTCGCAATATCAACTGAGGATTTGCTGATCGCA 7427  
 QY 1620 TACCAATATTTTCAAGGACACAGTGAACAAGCGGTTGATPAAGCACTATGACAGATC 1679  
 DB 7428 TACCAATATTTTCAAGGACACAGTGAACAAGCGGTTGATPAAGCACTATGACAGATC 7487  
 QY 1680 GTCAAAACTCTTCTGTGATGACTGAACGAAATTTCTGTAGAACCATATCTCCGCTTGG 1739  
 DB 7488 GTCAAAACTCTTCTGTGATGACTGAACGAAATTTCTGTAGAACCATATCTCCGCTTGG 7547  
 QY 1740 TGTCTCATGAGCGGATTCGCAAAAGCATGTATTTGCTGAGTATGATGTCGAGATCC 1799  
 DB 7548 TGTCTCATGAGCGGATTCGCAAAAGCATGTATTTGCTGAGTATGATGTCGAGATCC 7607  
 QY 1800 GAATGTAAAGTGTAGTGTGATCATGATTAACATGAGAGAGACCTCCACAAGCGTGT 1859  
 DB 7608 GAATGTAAAGTGTAGTGTGATCATGATTAACATGAGAGAGACCTCCACAAGCGTGT 7667  
 QY 1860 GGGCGTCCCGACAAAAAAGCATTAATGTCTGCTGCGGACGCAATTTGAGATCTTT 1919  
 DB 7668 GGGCGTCCCGACAAAAAAGCATTAATGTCTGCTGCGGACGCAATTTGAGATCTTT 7727  
 QY 1920 CCGGCGCTTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTC 1979  
 DB 7728 CCGGCGCTTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTC 7787  
 QY 1980 ACATGATTGGCTTACAGACGAATGCGGGGAGCTTTCAAACTCAACGGCGCTGTGA 2039

DB 7788 ACATGATTGGCTTACAGACGAATGCGGGGAGCTTTCAAACTCAACGGCGTGTGA 7847  
 QY 2040 GATATTTTATTTGAAGAACTTTCTTTCAAGCATGACACGGCTAATGATACCGGAGT 2099  
 DB 7848 GATATTTTATTTGAAGAACTTTCTTTCAAGCATGACACGGCTAATGATACCGGAGT 7907  
 QY 2100 TTAATTTGGGGGTGAGTGTCTTCTTCAAGTGTATGGGTGAGGAGGCTTATTCAGAC 2159  
 DB 7908 TTAATTTGGGGGTGAGTGTCTTCTTCAAGTGTATGGGTGAGGAGGCTTATTCAGAC 7967  
 QY 2160 CCGGTGAACGCGCTGTGCAATTAATCAATTTGAGAGCATTTTGGCAAGGCA 2219  
 DB 7968 CCGGTGAACGCGCTGTGCAATTAATCAATTTGAGAGCATTTTGGCAAGGCA 8027  
 QY 2220 TCTCTGAAACACTCTTGAAGAGATATACTACGCGAGTAATAATTAG 2268  
 DB 8028 TCTCTGAAACACTCTTGAAGAGATATACTACGCGAGTAATAATTAG 8076

RESULT 3  
 AAN50226  
 ID AAN50226 standard; DNA; 24593 BP.  
 XX  
 AC AAN50226;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 16-OCT-1991 (first entry)  
 XX  
 DE Sequence of opine synthase gene.  
 XX  
 KM Plant vector; Ti plasmid; T-DNA; se.  
 XX  
 OS Agrobacterium tumefaciens Ti plasmid.  
 XX  
 PN EP140556-A.  
 XX  
 PD 08-MAY-1985.  
 XX  
 PF 12-SEP-1984; 84EP-0306233.  
 XX  
 PR 14-SEP-1983; 83US-0532280.  
 XX  
 PA (AGRK ) AGRIGENETICS RES ASSOC LTD.  
 PA (LUBR ) LUBRIZOL GENETICS INC.  
 XX  
 PI Dahl GA, Sutton DW, Barker RF;  
 XX  
 DR WPI; 1985-112088/19.  
 XX  
 PT Plasmid contg. opine synthase gene for selection - and foreign  
 PT DNA, useful as vector for transforming plant cells  
 XX  
 PS Disclosure; Page 212-217; 69pp; Japanese.  
 XX  
 CC Plant cells (and protoplasts) and plasmids contg. the DNA fragment  
 CC which includes an opine synthase gene plus a gene for antibiotic  
 CC resistance are claimed. These plasmids provide max. efficiency for  
 CC transfer of foreign genes and can be amplified in the plant genome.  
 CC They do not contain genes specifying tumour formation and will not  
 CC spread antibiotic resistance throughout the plant population.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX  
 SQ Sequence 24593 BP; 6554 A; 5493 C; 5787 G; 6759 T; 0 other;  
 XX

Query Match 98.1%; Score 2225.8; DB 6; Length 24593;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 2253; Conservative 0; Mismatches 12; Indels 2; Gaps 2;  
 QY 1 ATGTGAGCTTCACTCTCTTGAATTAACAGTGCATCATCTCCAAACCAAAATGTGTGAT 60  
 DB 5809 ATGTGAGCTTCACTCTCTTGAATTAACAGTGCATCATCTCCAAACCAAAATGTGTGAT 5868

QY 61 CTGCAATGTCGATTAAGCGGATGAATTGAGCCGAGGGTTTCGATGCTCTTAGAA 120  
DB 5869 CTGCAATGTCGATTAAGCGGATGAATTGAGCCGAGGGTTTCGATGCTCTTAGAA 5928  
QY 121 CGAAGACCTTCTAGGGAGAGGATTACTCAAAATCTCCAGAGTGCAGGCTGGGTTA 180  
DB 5929 CGAAGACCTTCTAGGGAGAGGATTACTCAAAATCTCCAGAGTGCAGGCTGGGTTA 5988  
QY 181 GCTTGGCAAAAGGCTGGCGATGTCCTTCCCGAGATCTCAGCTGGTGGAAAGTACGA 240  
DB 5989 GCTTGGCAAAAGGCTGGCGATGTCCTTCCCGAGATCTCAGCTGGTGGAAAGTACGA 6048  
QY 241 GTTCTCTCCGCTTATCTATATTTGGCAAGAAATTCGGGGGGGATCTTGAATTCGAAA 300  
DB 6049 GTTCTCTCCGCTTATCTATATTTGGCAAGAAATTCGGGGGGGATCTTGAATTCGAAA 6108  
QY 301 CTTTGGGCGGGGCAAGTAGTAGTCTCGTTGCAATTCGACTTGGACACCATTTGGATG 360  
DB 6109 CTTTGGGCGGGGCAAGTAGTAGTCTCGTTGCAATTCGACTTGGACACCATTTGGATG 6168  
QY 361 GATTTCCTCGAAGCAACTAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA 420  
DB 6169 GATTTCCTCGAAGCAACTAATCCAAAGCCCTGTTTTGCTGAGCGGTAAAGATGTGCA 6228  
QY 421 CCGATGATCTAGTCAATTCGTCGGCAATTCCTAAGACTGCGCGCTTTCGAAAC 480  
DB 6229 CCGATGATCTAGTCAATTCGTCGGCAATTCCTAAGACTGCGCGCTTTCGAAAC 6288  
QY 481 CTGCGCAATGCGCTGTACGAGAAATGCGACGATGAAATGCGTTACCGGGTTTACCATACC 540  
DB 6289 CTGCGCAATGCGCTGTACGAGAAATGCGACGATGAAATGCGTTACCGGGTTTACCATACC 6348  
QY 541 CTTGAAAGGGGCGTGCATTTGACATGCTAGCTTATGCTGAAACCTGATGCTGAAGGGT 600  
DB 6349 CTTGAAAGGGGCGTGCATTTGACATGCTAGCTTATGCTGAAACCTGATGCTGAAGGGT 6408  
QY 601 TCGGCAAGGTTCTTTCCAAACAATGCACTGCTACGACGTCGACGACCGTTTTTTGACCAA 660  
DB 6409 TCGGCAAGGTTCTTTCCAAACAATGCACTGCTACGACGTCGACGACCGTTTTTTGACCAA 6468  
QY 661 TGTTCGATAGTGAAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGGAAAGTGGCG 720  
DB 6469 TGTTCGATAGTGAAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGGAAAGTGGCG 6528  
QY 721 GTCAATGGCGTGCATTTCCGACCTGTCGTGTGCAACGAACTGCTTCAATGCTGGGGTA 780  
DB 6529 GTCAATGGCGTGCATTTCCGACCTGTCGTGTGCAACGAACTGCTTCAATGCTGGGGTA 6588  
QY 781 GAGCATGTTACAAATATGAAAGCAAGTATGCTGTGTGAGGCAAGCTTTGGTCAATGCT 840  
DB 6589 GAGCATGTTACAAATATGAAAGCAAGTATGCTGTGTGAGGCAAGCTTTGGTCAATGCT 6648  
QY 841 TTCAGGAGACCTCCTAGTGTGTCGTGGCGAAATGGGGGCGATGGCATTTCTTCGCTGCA 900  
DB 6649 TTCAGGAGACCTCCTAGTGTGTCGTGGCGAAATGGGGGCGATGGCATTTCTTCGCTGCA 6708  
QY 901 TTTCTGCTGTTTTTTCTTCTCGACGCTTACGGCTGTCTTTCGATGAGACCGTTCCCAAT 960  
DB 6709 TTTCTGCTGTTTTTTCTTCTCGACGCTTACGGCTGTCTTTCGATGAGACCGTTCCCAAT 6768  
QY 961 CCGGCGCACTTCACACTTACTTGTGTACCAAGGCGTCCAAATCATGTGAAAGCCGGG 1020  
DB 6769 CCGGCGCACTTCACACTTACTTGTGTACCAAGGCGTCCAAATCATGTGAAAGCCGGG 6828  
QY 1021 CAGCTGCGACCGAAGCTGTTCATCGGCTTACCAAGGTTGGCGGTCTTGAAGAC 1080  
DB 6829 CAGCTGCGACCGAAGCTGTTCATCGGCTTACCAAGGTTGGCGGTCTTGAAGAC 6888  
QY 1081 GGTTCATAGCGAGATATTTGTTGGCTTCCGCTGCTATTAATCTCAGGCGCTTGAAA 1140  
DB 6889 GGTTCATAGCGAGATATTTGTTGGCTTCCGCTGCTATTAATCTCAGGCGCTTGAAA 6948

QY 1141 TCAGACACACTAGTAGTGAGGCTCATGACTCTGCGCAAAATTTGGCTGAACCGTTTCGGGAG 1200  
DB 6949 TCAGACACACTAGTAGTGAGGCTCATGACTCTGCGCAAAATTTGGCTGAACCGTTTCGGGAG 7008  
QY 1201 GAGTCTCTCTCTTACAGGATAGAGAGATCTTTCGTGGGCAACACTCTCTGTGTGT-AA 1259  
DB 7009 GAGTCTCTCTCTTACAGGATAGAGAGATCTTTCGTGGGCAACACTCTCTGTGTGTGAA 7068  
QY 1260 ACATGAGTTTCTCATGATTTGGGACCTTATTAAGCTAATGGGAAATAGATCTGGCGG 1319  
DB 7069 ACATGAGTTTCTCATGATTTGGGACCTTATTAAGCTAATGGGAAATAGATCTGGCGG 7127  
QY 1320 GTTTGGTCCAGTTTTTGAAGCGGGTTTATGAGATCTCTCGCTGTGATCAACGGATA 1379  
DB 7128 GTTTGGTCCAGTTTTTGAAGCGGGTTTATGAGATCTCTCGCTGTGATCAACGGATA 7187  
QY 1380 TGAAGAAATCAGCGGATGTCCCTGAAGGAATCTCAGAACTTTCAGCTGATTCGATC 1439  
DB 7188 TGAAGAAATCAGCGGATGTCCCTGAAGGAATCTCAGAACTTTCAGCTGATTCGATC 7247  
QY 1440 TGAAGTGTTAACGGTGTCTGTGAGCCAGCGCATATGCGCATGTTCAAGTCAAGCGGAT 1499  
DB 7248 TGAAGTGTTAACGGTGTCTGTGAGCCAGCGCATATGCGCATGTTCAAGTCAAGCGGAT 7307  
QY 1500 TCAGAGGAAAAAGCAAAAAATAAGATTAAGGCTTAAGCGGGATATCTGAATTTATGA 1559  
DB 7308 TCAGAGGAAAAAGCAAAAAATAAGATTAAGGCTTAAGCGGGATATCTGAATTTATGA 7367  
QY 1560 TAAAGTGTGTGTACATCTGCACTGCAAAATATCCAATCAGGACTTGCCTGACATGCGA 1619  
DB 7368 TAAAGTGTGTGTACATCTGCACTGCAAAATATCCAATCAGGACTTGCCTGACATGCGA 7427  
QY 1620 TACCAATTTTTTCAAGCACCACTGAAACCAAGCGTTGATTAACGCCATATGACAGATC 1679  
DB 7428 TACCAATTTTTTCAAGCACCACTGAAACCAAGCGTTGATTAACGCCATATGACAGATC 7487  
QY 1680 GTCAAAACTCTTCTGATGCTGAAAGAAATTCGAGTATGACATATCTCCGCTTG 1739  
DB 7488 GTCAAAACTCTTCTGATGCTGAAAGAAATTCGAGTATGACATATCTCCGCTTG 7547  
QY 1740 TGTCTCATGAGACGGGATGCAAAAGCAGTGTATTCCTGACCTATGAGTTCGAGATCC 1799  
DB 7548 TGTCTCATGAGACGGGATGCAAAAGCAGTGTATTCCTGACCTATGAGTTCGAGATCC 7607  
QY 1800 GAATGTTAAAGCTCTAGTGTCTATCACTTATACATGGAAGAGACTCCCAAGCTGTT 1859  
DB 7608 GAATGTTAAAGCTCTAGTGTCTATCACTTATACATGGAAGAGACTCCCAAGCTGTT 7667  
QY 1860 GCGGCTCCCGCAAAAAAGAGGATTATGTCTGCGGAGCGCAATTCGAGATCTTT 1919  
DB 7668 GCGGCTCCCGCAAAAAAGAGGATTATGTCTGCGGAGCGCAATTCGAGATCTTT 7727  
QY 1920 CCGGCGGTTTCCAGACACTATTTCTGCTGCGCTGATTAAGCAACAAATGTTATTCA 1979  
DB 7728 CCGGCGGTTTCCAGACACTATTTCTGCTGCGCTGATTAAGCAACAAATGTTATTCA 7787  
QY 1980 ACATGATGCTTACAGACGAGATGCGGGGAGGCTTTCAACTCAACCGGCTGTGA 2039  
DB 7788 ACATGATGCTTACAGACGAGATGCGGGGAGGCTTTCAACTCAACCGGCTGTGA 7847  
QY 2040 GGAATTTTATTCGAAAGACTTTTCTTAAGACACTGACACGGCTAATGATACCGGAGT 2099  
DB 7848 GGAATTTTATTCGAAAGACTTTTCTTAAGACACTGACACGGCTAATGATACCGGAGT 7907  
QY 2100 TTAATTTGGGCGGTTGAGTGTCTTCAACAGGTGATGGGTGAGGCTGTTTACAGAC 2159  
DB 7908 TTAATTTGGGCGGTTGAGTGTCTTCAACAGGTGATGGGTGAGGCTGTTTACAGAC 7967  
QY 2160 CCGGTGTAACCGCTGTGTGCAATTAATCAACAATTTGAGAGGCAATTTGGCAAGGCGAA 2219  
DB 7968 CCGGTGTAACCGCTGTGTGCAATTAATCAACAATTTGAGAGGCAATTTGGCAAGGCGAA 8027  
QY 2220 TCTCTCGAAACACTTGTGAAGAGATATACTAACCGCAATTAAG 2268

Db 8028 TCCTCTGGAACCTCTTGGAAGATATTACTACCGCAGCTAGAAATTAG 8076

RESULT 4  
AAV69756  
ID AAV69756 standard; DNA; 2211 BP.  
XX AAV69756;  
AC  
DT 01-MAR-1999 (first entry)  
XX  
DE Trypophan oxygenase laam gene.  
XX  
KM Trypophan oxygenase; laam gene; auxin; indoleacetic acid;  
KW transgenic plant; seedless fruit; tomato; watermelon; cucumber; ds.  
XX  
OS Agrobacterium tumefaciens.  
XX  
FH Key Location/Qualifiers  
FT CDS 4..2205  
FT /-tag= a  
XX  
PD MO9849888-A1.  
XX 12-NOV-1998.  
XX  
XX 06-MAY-1998; 98WO-US09013.  
XX  
XX 06-MAY-1997; 97US-0045725.  
XX  
XX (UNIV ) UNIV KANSAS STATE RES FOUND.  
XX  
XX L1 Y;  
XX  
XX MPI; 1999-034673/03.  
XX  
XX P-PSDB; AAV61574.  
XX  
XX A new construct to express phytohormones in developing fruit -  
PT useful for, e.g. producing substantially seedless fruit from  
PT transgenic plants  
XX  
XX  
XX Example 1; Page 21-26; 49pp; English.  
XX  
XX This is the nucleotide sequence of the Agrobacterium tumefaciens  
CC laam gene that codes for tryptophan oxygenase (see AAV61574). This  
CC enzyme catalyzes the conversion of tryptophan to indoleacetic acid.  
CC in the biosynthetic pathway of the auxin indoleacetic acid. A  
CC claimed DNA construct comprises either an isopenicillin transferase  
CC (see AAV61575) or a tryptophan oxygenase encoding sequence, operably  
CC linked to an ovary or developing fruit-specific plant-expressible  
CC promoter (see AAV69755 and AAV69759). The construct is used to stably  
CC integrate enzymes involved in cytokinin or auxin biosynthesis into  
CC the plant genome to achieve a transgenic plant (preferably tomato,  
CC cucumber or watermelon) producing seedless fruit in the absence of  
CC pollination.  
XX  
XX  
XX Sequence 2211 BP; 528 A; 504 C; 595 G; 584 T; 0 other;  
SQ  
Query Match 95.4%; Score 2162.8; DB 20; Length 2211;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2192; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 64 ACAATGTCGATAGGCGGATGAATTGACCGCAGGGTTCGATGCTCTTCTTGAACGA 123  
DB 1 ACAATGTCGATAGGCGGATGAATTGACCGCAGGGTTCGATGCTCTTCTTGAACGA 60  
QY 124 GAAGCTCTTAGGGGAAGAAGATTACTCAAGCTCCACCGAGTCCAGCGCTGGTTAGCT 183  
DB 61 GAAGCTCTTAGGGGAAGAAGATTACTCAATCTCCACCGAGTCCAGCGCTGGTTAGCT 120  
QY 184 TGCAAAAGGCTGGCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGGTAGCAGTT 243

Db 121 TGCAAAAGGCTGGCGATGTCGCTTCCCGAGATCTCAGCTGTGGAAGGTAGCAGTT 180  
QY 244 CTCTCCGCTTATATCTATATTTGGCAAGAAATCTGGGCGGATTAATTGAATGAAACCT 303  
Db 181 CTCTCCGCTTATATCTATATTTGGCAAGAAATCTGGGCGGATTAATTGAATGAAACCT 240  
QY 304 TGGGCGGCGCAACAGTGAAGTGTCTGTTGCCATGCACTTGGAACCAATTCATGAT 363  
Db 241 TGGGCGGCGCAACAGTGAAGTGTCTGTTGCCATGCACTTGGAACCAATTCATGAT 300  
QY 364 TTCTCCGAAGCAACAATAATCCAGGCCCTGTTTCTGAGCGGTAAGAGATGCAACCG 423  
Db 301 TTCTCCGAAGCAACAATAATCCAGGCCCTGTTTCTGAGCGGTAAGAGATGCAACCG 360  
QY 424 ATTGATCTTAGTATTTTGGCCATTTTCAATCTCTAAGCTCCGCGCTTTCGAACCTTG 483  
Db 361 ATTGATCTTAGTATTTTGGCCATTTTCAATCTCTAAGCTCCGCGCTTTCGAACCTTG 420  
QY 484 CCAATGCCGCTGACAGAAATGACAGATGAATAATGCGTTACCGGTTTACATAACCTT 543  
Db 421 CCAATGCCGCTGACAGAAATGACAGATGAATAATGCGTTACCGGTTTACATAACCTT 480  
QY 544 GAAAGGCGCGTGCATTTGACATGATGCTTATGTCGAAACCTGATGCTGAAGGCTTG 603  
Db 481 GAAAGGCGCGTGCATTTGACATGATGCTTATGTCGAAACCTGATGCTGAAGGCTTG 540  
QY 604 GCAGGTTCTTTCCAAACAATCGACTGCTCTACGACTGCAAGCCGTTTTTGAACCAATG 663  
Db 541 GCAGGTTCTTTCCAAACAATCGACTGCTCTACGACTGCAAGCCGTTTTTGAACCAATG 600  
QY 664 TCCGATAGTGAAGATGCGCTCTTCCGAGAGATGTTCTTAAGCGGAAGTGGCGGTC 723  
Db 601 TCCGATAGTGAAGATGCGCTCTTCCGAGAGATGTTCTTAAGCGGAAGTGGCGGTC 660  
QY 724 ATTGGCGCTGCAATTTCCGAGCTGTTGTCGCAACGAACCTGCTTACGCTGGGATGAC 783  
Db 661 ATTGGCGCTGCAATTTCCGAGCTGTTGTCGCAACGAACCTGCTTACGCTGGGATGAC 720  
QY 784 GATGTTACAATATATGAAGCAATGATCGTGTGAGGCAAGCTTTGTCACATGCTTTC 843  
Db 721 GATGTTACAATATATGAAGCAATGATCGTGTGAGGCAAGCTTTGTCACATGCTTTC 780  
QY 844 AGGAGCGCTCTAGTGTGTCGTCGGAATGAGGCGGATGCGATTTCTCTGTCGATTC 903  
Db 781 AGGAGCGCTCTAGTGTGTCGTCGGAATGAGGCGGATGCGATTTCTCTGTCGATTC 840  
QY 904 TGCCTGTTTTTCTCTCGAGCGTTACCGGCTGCTTTCGATGAGCGGTTCCCAATCCC 963  
Db 841 TGCCTGTTTTTCTCTCGAGCGTTACCGGCTGCTTTCGATGAGCGGTTCCCAATCCC 900  
QY 964 GGCACAGTGCACCTTACTTGTGTCCTACCAAGGCGTCCCAATGTCATGTCGAGGCGG 1023  
Db 901 GGCACAGTGCACCTTACTTGTGTCCTACCAAGGCGTCCCAATGTCATGTCGAGGCGG 960  
QY 1024 CTGCAACCGAAGCTGTTCCATCGCTTTACCAAGGTTGGCGTCTCTTGAAGACGCT 1083  
Db 961 CTGCAACCGAAGCTGTTCCATCGCTTTACCAAGGTTGGCGTCTCTTGAAGACGCT 1020  
QY 1084 TTTCATAGCGAGATATTTGTTGGCTTGCCTGTGCTATTAATCAAGGCTTGAATCA 1143  
Db 1021 TTTCATAGCGAGATATTTGTTGGCTTGCCTGTGCTATTAATCAAGGCTTGAATCA 1080  
QY 1144 GGCACATTAGTGGGCTCATGCTCTGCGCAATTTGCTGTAACCGTTTCGGGAGGAG 1203  
Db 1081 GGCACATTAGTGGGCTCATGCTCTGCGCAATTTGCTGTAACCGTTTCGGGAGGAG 1140  
QY 1204 TCTTCCTCTCAGAGATAGAGAGATCTTCTGGGCAACATCTCTGCTGCT-AAAACA 1262  
Db 1141 TCTTCCTCTCAGAGATAGAGAGATCTTCTGGGCAACATCTCTGCTGCTGTAACA 1200  
QY 1263 TGAAGTTTCTCTCATGATTTGGGACCTATTCAGCTATGGAATAGGATCTGGCGGGGTT 1322  
Db 1201 TGAAGTTTCTCTCATGATTTGGGACCTATTCAGCTATGGAATAGGATCTGGC-GGGTT 1259

QY 1323 TGGTCCAGTTTGTGAAGCGGGTTTATGAGATCTCCGGTGGTCTCAACGGATATGA 1382  
 DB 1260 TGGTCCAGTTTGTGAAGCGGGTTTATGAGATCTCCGGTGGTCTCAACGGATATGA 1319  
 QY 1383 AGAAATCAGCGGATGTCCTGGAAGAAATCTCAAACTTCCAGTCCGATCTGA 1442  
 DB 1320 AGAAATCAGCGGATGTCCTGGAAGAAATCTCAAACTTCCAGTCCGATCTGA 1379  
 QY 1443 AGTGGTTAAGCGGTGTCTGAGCGCGCATATGCGCATTTCAATCAGGGCGATTC 1502  
 DB 1380 AGTGGTTAAGCGGTGTCTGAGCGCGCATATGCGCATTTCAATCAGGGCGATTC 1439  
 QY 1503 GAAGAAAAACAAAAATTAAGATTAAGCTTAAAGAGGGGATTAATCTGAATTA 1562  
 DB 1440 GAAGAAAAACAAAAATTAAGATTAAGCTTAAAGAGGGGATTAATCTGAATTA 1499  
 QY 1563 GGTGGTGTCACTGTGAGTCTGCAATATCAACTCAGGCATTTGCTGACATGCGAT 1622  
 DB 1500 GGTGGTGTCACTGTGAGTCTGCAATATCAACTCAGGCATTTGCTGACATGCGAT 1559  
 QY 1623 CAATATTTTTCAGGCACCACTGAACCAAGGGTTGATTAACGCATATGACAGATCTC 1682  
 DB 1560 CAATATTTTTCAGGCACCACTGAACCAAGGGTTGATTAACGCATATGACAGATCTC 1619  
 QY 1683 AAAACTCTTCTGATGATGATGCAAAATTTCTGGTTAGACATATCTCCGCTTGTGT 1742  
 DB 1620 AAAACTCTTCTGATGATGATGCAAAATTTCTGGTTAGACATATCTCCGCTTGTGT 1679  
 QY 1743 CCTCATGAGCGGATGCAAAAGAGTATGCTGAGTATGAGTATGAGATCCGAA 1802  
 DB 1680 CCTCATGAGCGGATGCAAAAGAGTATGCTGAGTATGAGTATGAGATCCGAA 1739  
 QY 1803 TGGTAAAGGCTTGTGTCTCATGATTAATCATGGAGAGAGACTCCCAAGCTGTGGC 1862  
 DB 1740 TGGTAAAGGCTTGTGTCTCATGATTAATCATGGAGAGAGACTCCCAAGCTGTGGC 1799  
 QY 1863 GGTCCCGACAAAAGAGGATATGCTGCTGCGGAGCGCAATTTGAGATCTTTCC 1922  
 DB 1800 GGTCCCGACAAAAGAGGATATGCTGCTGCGGAGCGCAATTTGAGATCTTTCC 1859  
 QY 1923 GGGGTTTCCGACACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTCA 1982  
 DB 1860 GGGGTTTCCGACACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTCA 1919  
 QY 1983 TGAATTGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACCGCGTGTGAGA 2042  
 DB 1920 TGAATTGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACCGCGTGTGAGA 1979  
 QY 2043 TTTTATTTGGAAGACTTTTCTTCAAGACCTGACACGGCTAATGATCCGGAGTTTA 2102  
 DB 1980 TTTTATTTGGAAGACTTTTCTTCAAGACCTGACACGGCTAATGATCCGGAGTTTA 2039  
 QY 2103 CTGGGCGGGTGTGAGTGTCTTCAAGGTGATGGTGAAGGGTCTATTGACACCG 2162  
 DB 2040 CTGGGCGGGTGTGAGTGTCTTCAAGGTGATGGTGAAGGGTCTATTGACACCG 2099  
 QY 2163 GTGTAAAGCGGTGTGTGCAATTAATCAAAATGTGAGGCAATTTGGCAAAAGGCAATCC 2222  
 DB 2100 GTGTAAAGCGGTGTGTGCAATTAATCAAAATGTGAGGCAATTTGGCAAAAGGCAATCC 2159  
 QY 2223 TCTCGAAGCTCTTGAAGAGATTAATCAACCGCAGTAAGAAATTA 2268  
 DB 2160 TCTCGAAGCTCTTGAAGAGATTAATCAACCGCAGTAAGAAATTA 2205

RESULT 5  
 AAT48868/c  
 ID AAT48868 standard; DNA; 6112 BP.  
 AC AAT48868;  
 XX  
 XX  
 DT 07-APR-1997 (first entry)

XX Iams expression vector pPH5838.  
 DE Male, sterile plant; hybrid seed; pollen; microspore; oilseed;  
 XX  
 KW canola; rape; antisense RNA; stress resistance;  
 KW herbicide resistance; transgenic plant; crop protection;  
 KW Iams; indole acetamide synthase; Bp10; promoter; Brassica;  
 KW pPH5838; ds; cyclic.  
 XX  
 OS Chimeric Brassica sp.;  
 OS Chimeric Agrobacterium tumefaciens.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_difference 3794  
 FT /tag= a  
 FT /note= "base 3794 is given as v in the  
 specification"  
 PD MO9640950-AI.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; 96WO-US08692.  
 PR 07-JUN-1995; 95US-0476864.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PI Albani DJ, Arnison PG, Fabijanski SF, Laurian RS;  
 DR WPI; 1997-052340/05.  
 XX  
 PT Recombinant molecule for preparation of male sterile plants - used  
 PT to produce stress resistant plants and for increase in seed yield  
 XX  
 PS Example 39a; Fig 27a-e; 224pp; English.  
 XX  
 CC Plasmid pPH5838 (AAT48868) contains the Brassica microspore-specific  
 CC Bp10 promoter linked to the indole acetamide synthase (Iams) gene  
 CC from the Agrobacterium tumefaciens Ti plasmid derivative pPCV311,  
 CC and to the nos 3' non-translated region. The Iams gene is involved  
 CC in tumour formation in Agrobacterium-infected plants. Plasmids  
 CC pPH5838-41 (see also AAT48869-71), contg. Iams or Iams genes under  
 CC control of anther-specific Bp10 or TA39 promoters, were introduced  
 CC into canola lines by transformation. Male sterile plants were  
 CC obtd. These can be used in the prodn. of hybrid seed using pollen  
 CC from a male fertile line.  
 XX  
 SQ Sequence 6112 BP; 1515 A; 1466 C; 1451 G; 1679 T; 1 other;  
 Query Match 90.0%; Score 2040.8; DB 18; Length 6112;  
 Best Local Similarity 94.3%; Pred. No. 0;  
 Matches 2139; Conservative 0; Mismatches 127; Indels 2; Gaps 2;  
 QY 1 ATGTGAGCTTCACTCTCTGATTAACGAGTGCATCTCCCAACCAAAATGGTGAT 60  
 DB 2870 ATGTGAGCTTCACTCTCTGATTAACGAGTGCATCTCTCAACCAAAATGGTGAT 2811  
 QY 61 CTGACAAATGATGATTAAGCGGATGAATGACACCGAGGGTTCCGATCCTTTAGAA 120  
 DB 2810 CTGATATGATGATTAAGCGGATGAATGACACCGAGGGTTCCGATCCTTTAGAA 2751  
 QY 121 CGAAGAGCTTCTAGGGGAAGAGATTAATCAAGCTCCACCGAGTGCAGCGCTGGTTA 180  
 DB 2750 CGTGAAGCTTCTAGGGGAAGAGATTAATCAAGCTCCGAGAGTGCAGCGCTGGTTA 2691  
 QY 181 GCTTGCAAAAGGCTGGCGATGATGCTCCCGGAGATCTCAGCTGTGGAAGATGAGA 240  
 DB 2690 GCTTGCAAAAGGCTGGCGATGATGCTCCCGGAGATCTCAGCTGTGGAAGATGAGA 2631  
 QY 241 GTTCTCTCCGCTTATATCTATATGCAAGAAATTTCTGGGCGGATCTTGATGAAA 300  
 DB 2630 GCGCTCTCCGCTTATATCTATATGCAAGAAATTTCTGGGCGGATCTTGATGAAA 2571

QY 301 CCTTGGCGCGGCAACAGTAGTGTCTCGTTGCCATTCAGATTGACACCATTTGATG 360  
 Db 2570 CTTTGGCGCGGAGAAAGTAGTGTCTCGTTGCCATTCAGATTGATG 2511  
 QY 361 GATTTCCTCGAAGCAACTATTCGAACCCCTGTTTGTGAGCGGTAAGATGTGA 420  
 Db 2510 GATTTCCTCGAAGCAACTATTCGAACCCCTGTTTGTGAGCGGTAAGATGTGA 2451  
 QY 421 CCGATTGATCTTGTAGTGTCTCGTTGCCATTCAGATTGATG 480  
 Db 2450 TCCAGCATCTTGTAGTGTCTCGTTGCCATTCAGATTGATG 2291  
 QY 481 CTGCGCAATGCGGCTGTAGAGAAATGCGATGAAATGCGTTACCGATTACCA 540  
 Db 2390 CTGCGCAATGCGGCTGTAGAGAAATGCGATGAAATGCGTTACCGATTACCA 2331  
 QY 541 CTGCGCAATGCGGCTGTAGAGAAATGCGATGAAATGCGTTACCGATTACCA 600  
 Db 2330 CTGCGCAATGCGGCTGTAGAGAAATGCGATGAAATGCGTTACCGATTACCA 2271  
 QY 601 TCGGCAAGGCTCTTTCGAACATGAGTGTCTGATGATGATGATGATGATGATG 660  
 Db 2270 TCGGCAAGGCTCTTTCGAACATGAGTGTCTGATGATGATGATGATGATGATG 2211  
 QY 661 TGTTCGATAGTGAAGGATGCGCTCTTTCGAGAGATGTTCTTAAGCGAAAGTGG 720  
 Db 2210 TGTTCGATAGTGAAGGATGCGCTCTTTCGAGAGATGTTCTTAAGCGAAAGTGG 2151  
 QY 721 GTGATGCGGCTGTAGATTTCCGAGCTGTGTGTGCAACGAACTGCTTCAATGCTGG 780  
 Db 2150 ATGATGCGGCTGTAGATTTCCGAGCTGTGTGTGCAACGAACTGCTTCAATGCTGG 2091  
 QY 781 GAGCATTTCAATATATGAAAGCAAGATGCTGTGTGAGGCAAGCTTGTGATCATGCT 840  
 Db 2090 GAGCATTTCAATATATGAAAGCAAGATGCTGTGTGAGGCAAGCTTGTGATCATGCT 2031  
 QY 841 TTCAGGAGCGCTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900  
 Db 2030 TTCAGGAGCGCTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1971  
 QY 901 TTTGCTGTTGTTTCTTCTGTGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
 Db 1970 TTTGCTGTTGTTTCTTCTGTGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1911  
 QY 961 CCGGCGCAAGTCACTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020  
 Db 1910 CCGGCGCAAGTCACTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1851  
 QY 1021 CAGGTGCGACGAAAGTGTTCATCGGTTTCAACAGGTTGCGGTCTTGTGAAGGAC 1080  
 Db 1850 CAGGTGCGACGAAAGTGTTCATCGGTTTCAACAGGTTGCGGTCTTGTGAAGGAC 1791  
 QY 1081 GGTTCATGAGCAGATATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140  
 Db 1790 GGTTCATGAGCAGATATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1731  
 QY 1141 TCAGGACACATTTAGT 1200  
 Db 1730 TCAGGACACATTTAGT 1671  
 QY 1201 GAGTCTTCTCTTCAAGGATATGAGAGATCTTCTGTGAGCAGATCTCTCGTGTGTGT 1259  
 Db 1670 GAGTCTTCTCTTCAAGGATATGAGAGATCTTCTGTGAGCAGATCTCTCGTGTGTGT 1611  
 QY 1260 ACATGAGTTTCTCTCATGATGAGACCTATTCAGGTAATGGAATAGGATCTGTGGCG 1319  
 Db 1610 ACATGAGTTTCTCTCATGATGAGACCTATTCAGGTAATGGAATAGGATCTGTGGCG 1552  
 QY 1320 GTTGTGTCAGTTTGTGAAGCGGTTTATTTGAATCTCTCGTGTGTGTGTGTGTGTGTGT 1379  
 Db 1551 GTTGTGTCAGTTTGTGAAGCGGTTTATTTGAATCTCTCGTGTGTGTGTGTGTGTGTGT 1492

QY 1380 TGAAGAAAAACAGCGGATGTGCTTGAAGAAATCTGAGAACTTTCAGATGTCGATC 1439  
 Db 1491 TGAAGAAAAACAGCGGATGTGCTTGAAGAAATCTGAGAACTTTCAGATGTCGATC 1432  
 QY 1440 TGAAGTTGTTTAAAGGT 1499  
 Db 1431 TGAAGTTGTTTAAAGGT 1372  
 QY 1500 TCAGAGAAAAACAAAAATTAAGATTAAGGCTTAAAGCGGATATCTGAATTTATGA 1559  
 Db 1371 TGAAGAAAAACAAAAATTAAGATTAAGGCTTAAAGCGGATATCTGAATTTATGA 1312  
 QY 1560 TAAAGT 1619  
 Db 1311 TAAAGT 1252  
 QY 1620 TACCAATATTTTCAAGGACCAAGTGAACCAAGGTTGATTAACGCCATTTAGCAGATC 1679  
 Db 1251 TACCAATATTTTCAAGGACCAAGTGAACCAAGGTTGATTAACGCCATTTAGCAGATC 1192  
 QY 1680 GTCAAAATCTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1739  
 Db 1191 GTCAAAATCTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1132  
 QY 1740 TGTCTCATGAGAGGATGCGAAAGCAGTATGCTGTGATGATGATGATGATGATGATG 1799  
 Db 1131 TGTCTCATGAGAGGATGCGAAAGCAGTATGCTGTGATGATGATGATGATGATGATG 1072  
 QY 1800 GAATGTTAAAGTCTAGTGTCTCATGATTTATATGAGAGAGAGATCTCCCAAGCTGT 1859  
 Db 1071 GAATGTTAAAGTCTAGTGTCTCATGATTTATATGAGAGAGAGATCTCCCAAGCTGT 1012  
 QY 1860 GGGGATCCCCGACAAAAAGAGGATTTATGTCTGTGCGGAGCGCAATTTGAGATCTTT 1919  
 Db 1011 GGGGATCCCCGACAAAAAGAGGATTTATGTCTGTGCGGAGCGCAATTTGAGATCTTT 952  
 QY 1920 CCGGCGGTTTCCGACACCTATTTCTGTGCTGTGCTGTGATTTACGACCAAAATGTTTCA 1979  
 Db 951 CCGGCGGTTTCCGACACCTATTTCTGTGCTGTGCTGTGATTTACGACCAAAATGTTTCA 892  
 QY 1980 ACATGATGCTTACAGAGAAATGCGGGGAGCTTTCAAACTCAACCGGCTGTGTGA 2039  
 Db 891 ACATGATGCTTACAGAGAAATGCGGGGAGCTTTCAAACTCAACCGGCTGTGTGA 832  
 QY 2040 GGAATTTTATTTGAGAACTTTCTTTCAAGCACTGACACCGGCTAATGATACCGGAGT 2099  
 Db 831 GGAATTTTATTTGAGAACTTTCTTTCAAGCACTGACACCGGCTAATGATACCGGAGT 772  
 QY 2100 TTAATTGCGGAGTTGAGATTTCTTCAACAGGTGATGAGGTGATGATGATGATGATG 2159  
 Db 771 TTAATTGCGGAGTTGAGATTTCTTCAACAGGTGATGAGGTGATGATGATGATGATG 712  
 QY 2160 CCGGTGTACGCGGT 2219  
 Db 711 CCGGTGTACGCGGT 652  
 QY 2220 TCCCTGGAACACTTTGGAAGATATTAATCAACCGAGTAAATTA 2267  
 Db 651 TCCCTGGAACACTTTGGAAGATATTAATCAACCGAGTAAATTA 604

RESULT 6  
 AAT4870/c  
 ID AAT4870 standard; DNA; 7208 BP.  
 XX  
 AC AAT4870;  
 XX  
 DT 07-APR-1997 (first entry)  
 XX  
 DE 1amh expression vector pHP5840.  
 XX  
 KM Male sterile plant; hybrid seed; pollen; microspore; oilseed;  
 KM canola; rape; antisense RNA; stress resistance;

QY	361	GATTCTCCGAAGACAACAATTACCAAGCCCTGTTTTTGGTGAAGCCGTTAAAGATGAGCA	420
Db	2564	GATTTCTCCGAAGACAACAATTCTCAAAACCTGTTTTTGGTGAAGCCGTTAAAGATGAGCA	2505
QY	421	CCGATGTACTTAACTCAATTCGTCGTCGTCATCTCTAAGACTGCGCGCTTTCGAAAC	480
Db	2504	TCGACGAGATCTTAACTCAATTCGTCGTCATCTCTAAGACTGCGCGCTTTCGAAAC	2445
QY	481	CTGCCAATGCCGCTGTACGAGAATGGACGATGAATGCGTTACCGGCTTTCATTAAC	540
Db	2444	CTGCAAAATGCCGCTTTCGAGAAAGGACGACGCAAAACGCGTTTACCGGCTTTCATTAAC	2385
QY	541	CTTGAAGGGGCGTGCATTTGACATGTATGCTTATGTTGAAACCGTATGCTGAAGGT	600
Db	2384	CTTGAAGGGGCGTGCATTTGACATGTATGCTTATGTTGAAACCGTATGCTGAAGGT	2325
QY	601	TCGGCAGGTTCTTTCCAAACATCGACTGTCTACACATGACAGACGTTTTCGACAA	660
Db	2234	TCGGCAGGTTCTTTCCAAACATCGACTGTCTACACATGACAGATCGTTTTCGACAA	2265
QY	661	TGTTCCGATATGTGACGGAATGCGCTCTTTCCGAGGATGTTCTTAACCGAAAGTGGC	720
Db	2264	TGTTCCGATATGTGACGGAATGCGCTCTTTCCGAGGATGTTCTTAACCGAAAGTGGC	2205
QY	721	GTCATTTGGCGGTGGCATTTCCGGACTGTGTGTGGCAACGAACTGCTCAATGCTGGGGTA	780
Db	2204	ATCATATGCGGTGGCATTTCCGGACTGTGTGTGGCAACGAACTGCTCAATGCTGGGTGA	2145
QY	781	GACGATGTTACAATATATGAAGCAAGATCGTGTGGAAGGCAAGCTTTGGTCAATGCT	840
Db	2144	GACGATGTTACAATATATGAAGCAAGATCGGTTTGAAGGCAAGCTTTGGTCAATGCT	2085
QY	841	TTCAAGGACGCTCTTAAGTGTGTGGCCGAAATGGGGGCGATGCGATTTCTCTGTCTGA	900
Db	2084	TTCAAGGACGCTCTCAACGTGTGTGGCCGAAATGGGGGCGATGCGATTTCTCTGTCTGA	2025
QY	901	TTTCGCTGTGTTTTCTTCCTCGAGCGTTAGGGCGTCTCGAATGAGGCGGTTCCCAAT	960
Db	2024	TCGTGCTGTGTTTTCTTCCTCGAGCGTTAGGGCGTCTTCGATGAGGCGGTTCCCAAT	1965
QY	961	CCCGGCAACAGTCGACATTAATGTTGTTACCAAGCGCTCCAAATACATGTGGAAGCCGG	1020
Db	1964	CCCGGCAACAGTCGACATTAATGTTGTTACCAAGCGCTCCGATATATGTGGAAGCCGG	1905
QY	1021	CAGCTGCCACCGAAGCTGTTCATCGCGTTTACAACGTTGGCGTTCCTTGAAAGAC	1080
Db	1904	CAGCAGCACCGAAGCGTTCATCGCGTTTACAACGCGTTGGCGTTCCTTGAAAGAC	1845
QY	1081	GGTTTTCATGAGGAGATATTTGTTGGGCTTCCGCTGTGCTATTAATCTCAGGCGTTGAA	1140
Db	1844	GGTTTTCATGAGGAGATATTTGTTGGGCTTCCGCTGTGCTATTAATCTCAGGCGTTGAA	1785
QY	1141	TCAGGACACATTAAGTGGGCTCACTGCTGTGCAAAATTTGCTGAACCGTTTTCGGAGG	1200
Db	1784	TCAGGACACATTAAGTGGGCGCTCACTGCTGTGCAAAATTTGCTGAACCGTTTTCGGAGG	1725
QY	1201	GAGTCTTCTCTTCAGGGATAGAGAAGATCTTTCTGGGCAACATCTCTCTGGTGT-AA	1259
Db	1724	GAGTCTTCTCTTCAGGGATAGAGAAGATCTTTCTGGGCAACGATCTCTCTGGTGTGAA	1665
QY	1260	ACATGAGGTTTCTCTCAATGATTTGGGACCTAATTCAAGCTAATGGGAATAGGATCTGGCGG	1319
Db	1664	ACATGAGGTTTCTCTCAATGATTTGGGACCTAATTCAAGCTAATGGGAATAGGATCTGGC-GG	1606
QY	1320	GTTTGTGTCAGTTTTCGAAAGCGGTTTATGTAGATCTTCGCTGTGTCAATCAACGATA	1379
Db	1605	GTTTGTGTCAGTTTTCGAAAGCGGTTTATGTAGATCTTCGCTGTGTCAATCAACGATA	1546
QY	1380	TGAAGAAATTCACCGGATGTGCGCTTAAGGAATCTCAAGACTTCCACGTGCGATTCGATC	1439
Db	1545	TGAAGAAATTCACCGGATGTGCTTAAGGAATCTCAAGACTTCCACGTGCGAATTCACAC	1486

QY	1440	TGAAGTGGTTAACGGTGTGTGTGGACGACGCAATATGCAATGTTCACAGTCAGGCGCAT	1499
Db	1485	TCAAGTGGTTAACGGTGTGTGTGGACGACGCAATATGCAATGTTCACAGTCAGGCGCAT	1426
QY	1500	TCAGAAAGAAAAGACAAAATAAAGATTAAGCTTAAGACGCGGATATCTGAATCTTATGA	1559
Db	1425	TGAGAAAGAAAAGACAAAATAAAGATTAAGCTTAAGACGCGGATATCTGAATCTTATGA	1366
QY	1560	TAAAGTGGTGTCACTCTGGAATTCGCAAAATATCCAACTCAGGCAATTGCCTCGACATGGGA	1619
Db	1365	TAAAGTGGTGTTCATCTGGACCTCGCAAAATATCCAACTCAGGCAATTGTCTGCAATGGGA	1306
QY	1620	TACCAATATTTTTGAGCGACACAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATC	1679
Db	1305	TACCAACATTTTTGTGTGACCAAGTGAACCAAGCGGTGATTAACAGCCATATGACAGGCTC	1246
QY	1680	GTCAAAACCTTCTCTGATGACTGAACGAAAATTCGTGTAGACCAATATCTCCGCTTGG	1739
Db	1245	GTCAAAACCTCTCTGCTGACTGAACGAAAATTTGGTTAGACCAATATCTCCGCTCTG	1186
QY	1740	TGTCCTCATGGAACGGGATTCGAAAAGCAAGTATATTCCTGGAATATAGATGGACAGATCC	1799
Db	1185	TGTCCTCATGGAACGGGATTCGAAAAGCAAGTATATTCCTGGAATATAGAGCCACAGACCC	1126
QY	1800	GAATGCTAAAGGTGTAGTGTCTCATCAGTTATATACATGGAGAGCACTCCACAAGCTGT	1859
Db	1125	GAAATGCTAAAGGTGTGTGTCTCATCAGTTATATACATGGAGAGCACTCCACAAGCTGT	1066
QY	1860	GCGCGTCCCGACAAAAGAGCGATTAATGTCTGTCTGCGGACGCAATTTGAGATCTTT	1919
Db	1065	GCGCGTTCGCCACAAAAGAGCGATTTCTGTCTGTGCGGACGCAATTTGAGATCTTT	1006
QY	1920	CCGGGCGTTTGGCCAGACCTATTTCCGTGCGTGGCGCTATTAACGACCAAAATGTTATTTCA	1979
Db	1005	CCGGGCGTTTGGCCAGACCTATTTCCGTGCGTGGCGCTATTAACGACCAAAATGTTATTTCA	946
QY	1980	ACATGATTTGGCTTACAGACGAAATGCGGGGAGCTTTCAAATCAACCGCGTGTGA	2039
Db	945	ACATGATTTGGCTTACAGACGAAATGCGGGGAGCTTTCAAATCAACCGCGTGTGA	886
QY	2040	GGATTTTTTATTTCTGAAGACTTTTCTTTCAAGCACTGGAACCGGCTAATGATATCCGAGT	2099
Db	885	GGATTTTTTATTTCTGAAGACTTTTCTTTCAAGCGCTGGAACATGACTAATGATATCCGAGT	826
QY	2100	TTACTTTGGCGGGTTGCAGTGTTCCTTCAACAGTGAATGGGTGAGGGTGCATTTCAAC	2159
Db	825	TTACTTTGGCGGGTTGCAGTGTTCCTTCAACGGTGAATGGGTGAGGGTGCATTTCAAC	766
QY	2160	CGCGTGTAAACGCGTCTGTGCAATTTATCACAATTTGTGAGGCAATTTGGCAAAAGGCAA	2219
Db	765	CGCGTGTAAACGCGTCTGTGCAATTTATCACAATTTGTGAGGATTTTGGCAAAAGGCAA	706
QY	2220	TTCTCTCGAACAATCTTGGAGAGATATTAATCAACCGACAGTAAATAA 2267	
Db	705	TTCTCTCGAACAATCTTGGAGAGATATTAATCAACCGACATATGAATAA 658	

RESULT 7  
AAD00626  
ID AAD00626 standard; DNA; 2544 BP.

AC AAD00626;

DT 29-AUG-2000 (first entry)

DE Agrobacterium tumefaciens ipt/iaam untranslatable BR construct.

KW laam; laah; ipt; T-DNA; transferred DNA; gall-disease resistance; BR;

KW plant growth hormone; ds.

**OS Agrobacterium tumefaciens**

Feature	Location/Qualifiers
PH Key	1..309
FT misc_feature	/+tag= a
FT	/note= "Corresponds to Agrobacterium mutant ipt gene"
FT	303..2105
misc_feature	/+tag= b
FT	/note= "Corresponds to Agrobacterium mutant iaam gene"
FT	

PN	WO200026346-A1
XX	
PD	11-MAY-2000.

PF 04-NOV-1999; 99WO-US26100.

PR 05-NOV-1998; 98US-0107185.

PA (UYOR-) UNIV OREGON STATE.

PI Ream W, Mok MC, Lee H;

DR WPI; 2000-365599/31.

PT Gall resistant plants generated by transforming cells with an

10

PS Example c; Page 44-45; 49pp; English.

CC The patent discloses a method for producing plants resistant to gall

CC bacterial resistance (BR) construct capable of eliciting co-suppression.

CC homologous to the tumour or gall disease-causing genes *iaam*, *iaah* and

CC introduction of premature termination codons or frameshift mutations, to

CC of plants to gall diseases. The present DNA sequence is the *Agrobacterium*

CC RNA forms of both the mutant *ipt* and *iaam* genes. This construct can

[illegible][illegible]

Best Local Similarity 99.8%; Pred. No. 0;  
 Waterbed 1701; Consistency 0; Wicnatchob 4; Indol 0; Cass 0

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

318  
327

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 84

[illegible]

133 18

438 498

[illegible][illegible][illegible][illegible][illegible][illegible]

**A**

\_\_\_\_\_

Db 678 AGCAACAATAATCCAGCCCTGTTTGTGAGGCGGTAAAAATGTGACCCGATTTGATCT 737  
 Qy 432 TAGTCATTTGCTGGCCATTTTCAATCTCTAAGATGCGGGCTTTGGAAACCCGCAATGCC 491  
 Db 738 TAGTCATTTGCTGGCCATTTTCAATCTCTAAGATGCGGGCTTTGGAAACCCGCAATGCC 797  
 Qy 492 GCTGTACGAGATGAGCAAGATGAATGCGTTACCGGGTTTACATTAACCTTTGAAGGGGC 551  
 Db 798 GCTGTACGAGATGAGCAAGATGAATGCGTTACCGGGTTTACATTAACCTTTGAAGGGGC 857  
 Qy 552 CGTCCCATTTGACATGCTAGCTTATGCTCGAAACCTGATGCTGAAGGGTTGCGCAGGTT 611  
 Db 858 CGTCCCATTTGACATGCTAGCTTATGCTCGAAACCTGATGCTGAAGGGTTGCGCAGGTT 917  
 Qy 612 CTTTCCAAACATGCACTTGTCTACGACTGAGACCGTTTTTTTGAACCAATGTTCCGATAG 671  
 Db 918 CTTTCCAAACATGCACTTGTCTACGACTGAGACCGTTTTTTTGAACCAATGTTCCGATAG 977  
 Qy 672 TGAACGATCGGCTTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTCATTGGCGC 731  
 Db 978 TGAACGATCGGCTTCTTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGTCATTGGCGC 1037  
 Qy 732 TGGCATTTCCGAGATCGTGGTGGCAACGAACTGCTTATGCTGGGGTGAACGATGTTAC 791  
 Db 1038 TGGCATTTCCGAGATCGTGGTGGCAACGAACTGCTTATGCTGGGGTGAACGATGTTAC 1097  
 Qy 792 AATATGAGACGATGATGCTGTGAGGCAACCTTTGGCAATGCTTTGAGGAGCGC 851  
 Db 1098 AATATGAGACGATGATGCTGTGAGGCAACCTTTGGCAATGCTTTGAGGAGCGC 1157  
 Qy 852 TCCTAGTGTGCTGGCCGAAATGGGGGCGATGCGATTTCTCTGCTGATCTGCTGTT 911  
 Db 1158 TCCTAGTGTGCTGGCCGAAATGGGGGCGATGCGATTTCTCTGCTGATCTGCTGTT 1217  
 Qy 912 TTTTCTTCTGAGCGGTTACGCGCTGTCTTGATGAGCGCGTTCCAAATCCGCGACAGT 971  
 Db 1218 TTTTCTTCTGAGCGGTTACGCGCTGTCTTGATGAGCGCGTTCCAAATCCGCGACAGT 1277  
 Qy 972 CGAAGCTTACTTGGTCTACCAAGGCGTCCAAATATGATGGAAGACCGGCGACGCTGCCAC 1031  
 Db 1278 CGAAGCTTACTTGGTCTACCAAGGCGTCCAAATATGATGGAAGACCGGCGACGCTGCCAC 1337  
 Qy 1032 GAAGCTTCTCATGCGCTTTACCAAGGTTGGCGCTGACCGTTTGGGAGGAGTCTTCTC 1091  
 Db 1338 GAAGCTTCTCATGCGCTTTACCAAGGTTGGCGCTGACCGTTTGGGAGGAGTCTTCTC 1397  
 Qy 1092 GCGAGATATTGTGTGGCTTGCCTGTGCTATTACTCAGGCGCTTGAATCAGACACAT 1151  
 Db 1398 GCGAGATATTGTGTGGCTTGCCTGTGCTATTACTCAGGCGCTTGAATCAGACACAT 1457  
 Qy 1152 TAGGTGGCTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1211  
 Db 1458 TAGGTGGCTCATGATCTCTGCGCAAAATTTGGCTGAACCGTTTGGGAGGAGTCTTCTC 1517  
 Qy 1212 TTCAGGATAGAGAGATCTTCTTGGGACACATCCTCGTGGTGAACATGGAAGTTT 1271  
 Db 1518 TTCAGGATAGAGAGATCTTCTTGGGACACATCCTCGTGGTGAACATGGAAGTTT 1577  
 Qy 1272 CTTATGATTGGACCTATTCAAGCTATGGAATAGGATCTGGCGGGGTTTGGTCCAGT 1331  
 Db 1578 CTTATGATTGGACCTATTCAAGCTATGGAATAGGATCTGGCGGGGTTTGGTCCAGT 1637  
 Qy 1332 TTTTGAAGGGGTTTATTGAGATCTCTCGCTTGGTATCAAGCGATATGAAGAAATCA 1391  
 Db 1638 TTTTGAAGGGGTTTATTGAGATCTCTCGCTTGGTATCAAGCGATATGAAGAAATCA 1697  
 Qy 1392 GCGGATGCGCTGGAAGAAATCTCAGAACTTCCAGTCCGATCGCATCTGAAGGTTAA 1451  
 Db 1698 GCGGATGCGCTGGAAGAAATCTCAGAACTTCCAGTCCGATCTGAAGGTTAA 1757  
 Qy 1452 CGGTGTGTCTGTGAGCCAGCGCATATGCTTCAAGTCAAGGCGATTCAGAAAGAAA 1511  
 Db 1758 CGGTGTGTCTGTGAGCCAGCGCATATGCTTCAAGTCAAGGCGATTCAGAAAGAAA 1817

Qy 1512 GACAAAATTAAGATTAAGGCTTAAGACCGGATATCTGAACCTTTATGATTAAGGTGCT 1571  
 Db 1818 GACAAAATTAAGATTAAGGCTTAAGACCGGATATCTGAACCTTTATGATTAAGGTGCT 1877  
 Qy 1572 CACATCTGACCTGCAAAATATCCAACTCAGGACTTCCGATGCGATACCAATATTTT 1631  
 Db 1878 CACATCTGACCTGCAAAATATCCAACTCAGGACTTCCGATGCGATACCAATATTTT 1937  
 Qy 1632 TCAGGACACGATGAACCAAGCGGTTGATTAACAGCCATATGACAGATTCGCAAACTTT 1691  
 Db 1938 TCAGGACACGATGAACCAAGCGGTTGATTAACAGCCATATGACAGATTCGCAAACTTT 1997  
 Qy 1692 CTTGATGACTGAACGAAAATTTCTGTTAGACATATCCTCCGCTTGTGCTCATGGA 1751  
 Db 1998 CTTGATGACTGAACGAAAATTTCTGTTAGACATATCCTCCGCTTGTGCTCATGGA 2057  
 Qy 1752 CGGGATCGCAAAAGCAGTATTCCTGCACTATGAGTCGACAGATCCGAATGCT 1806  
 Db 2058 CGGGATCGCAAAAGCAGTATTCCTGCACTATGAGTCGACAGATCCCACTGCT 2112

RESULT 8  
 AAD00625  
 ID AAD00625 standard; DNA; 1807 BP.  
 XX  
 AC AAD00625;  
 XX  
 DT 29-AUG-2000 (first entry)  
 XX  
 DE Agrobacterium tumefaciens T-DNA mutant iaam gene.  
 XX  
 KM iaam; iaam; ipt; T-DNA; transferred DNA; gall-disease resistance; BR;  
 KM bacterial resistance construct; untranslatable RNA; tumour; oncogene;  
 KM auxin; plant growth hormone; mutant; ds.  
 OS Agrobacterium tumefaciens.  
 XX  
 PN W0200026346-A1.  
 XX  
 PD 11-MAY-2000.  
 XX  
 PF 04-NOV-1999; 99MO-US26100.  
 XX  
 PR 05-NOV-1998; 98US-0107185.  
 XX  
 PA (UYOR-) UNIV OREGON STATE.  
 XX  
 PI Ream W, Mok MC, Lee H;  
 XX  
 DR WPI; 2000-365599/31.  
 XX  
 PT Gall resistant plants generated by transforming cells with an  
 PT untranslatable nucleic acid homologous to a gall disease-causing gene  
 PT -  
 PS Example c; Page 43-44; 49pp; English.  
 XX  
 CC The patent discloses a method for producing plants resistant to gall  
 CC disease, induced by Agrobacterium infection, by transformation with a  
 CC bacterial resistance (BR) construct capable of eliciting co-suppression.  
 CC The construct encodes an untranslatable RNA molecule, that is highly  
 CC homologous to the tumour or gall disease-causing genes iaam, iaam and  
 CC ipt of Agrobacterium tumefaciens. These oncogenes are modified by  
 CC introduction of premature termination codons or frameshift mutations, to  
 CC inhibit their expression. This method is used to reduce susceptibility  
 CC of plants to gall diseases. The present DNA sequence is the Agrobacterium  
 CC tumefaciens T-DNA (transferred DNA) mutant iaam gene. This oncogene was  
 CC modified by introduction of a stop sequence (TGA) at the third codon and  
 CC deletion of two bases following the third codon. This sequence was used  
 CC for the construction of double stranded RNA and untranslatable dsRNA  
 CC vectors, that reduce gall formation in transformed plant cells.  
 XX

SQ Sequence 1807 BP; 428 A; 418 C; 486 G; 475 T; 0 other;  
 Query Match 78.8%; Score 1788; DB 21; Length 1807;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 12 ACCTCTCTGATTAACGATGCGATCATCTCCCAACCAAAATGATGATCTGAACATGCT 71
DB 18 ACCTCTCTGATTAACGATGCGATCATCTCCCAACCAAAATGATGATCTGAACATGCT 77
QY 72 CGATTAAGCGCGATTAATGGAACCGCAGGGTTCCGATGCCCTTTCTTAAGAAGAAAGCTTC 131
DB 78 CGATTAAGCGCGATTAATGGAACCGCAGGGTTCCGATGCCCTTTCTTAAGAAGAAAGCTTC 137
QY 132 TAGGGGAAGAGGATTAATCAAGCTCCACGAGTGCAGCGTGGGTAGCTTGCAAAAG 191
DB 138 TAGGGGAAGAGGATTAATCAAGCTCCACGAGTGCAGCGTGGGTAGCTTGCAAAAG 197
QY 192 GCTGGCCGATGTCGCTTCCCGAGATCTGAGCTGTGAAAGGTAGCAGTTCTCTCGC 251
DB 198 GCTGGCCGATGTCGCTTCCCGAGATCTGAGCTGTGAAAGGTAGCAGTTCTCTCGC 257
QY 252 TTATATCTATTAATGGCAAGAAATTTCTGGGCGGATCTTGAATCGAAACCTTGGGCGG 311
DB 258 TTATATCTATTAATGGCAAGAAATTTCTGGGCGGATCTTGAATCGAAACCTTGGGCGG 317
QY 312 GGCAACGATAGTGTCTCGTGCATCGAATTGGAACCACTTTGATGATTTCTCGA 371
DB 318 GGCAACGATAGTGTCTCGTGCATCGAATTGGAACCACTTTGATGATTTCTCGA 377
QY 372 AGCAACAATACTCAACAGCCCTGTTTTTCTGAGCGGTTAAAGATGTCACCGATGATCT 431
DB 378 AGCAACAATACTCAACAGCCCTGTTTTTCTGAGCGGTTAAAGATGTCACCGATGATCT 437
QY 432 TAGTCATTTGTGGCCATTTCATCTCTAAGATCTGCCGCTTTTGAAACCTTCCCAATGCC 491
DB 438 TAGTCATTTGTGGCCATTTCATCTCTAAGATCTGCCGCTTTTGAAACCTTCCCAATGCC 497
QY 492 GCTGTAGAGAAATGGCAAGTGAATGCGTTACCGGGTTTACCAATAACCTTGAAAGGGG 551
DB 498 GCTGTAGAGAAATGGCAAGTGAATGCGTTACCGGGTTTACCAATAACCTTGAAAGGGG 557
QY 552 CGTGCCATTTGACATGCTAGCTTAATGCTGCAACCTGATGCTGAAGGGTTCCGCAAGTTC 611
DB 558 CGTGCCATTTGACATGCTAGCTTAATGCTGCAACCTGATGCTGAAGGGTTCCGCAAGTTC 617
QY 612 CTTTCCAAACAATCGACTTGTCTTAAGACTGCAACCGTTTTTTTGAACCAATGTTCCGATAG 671
DB 618 CTTTCCAAACAATCGACTTGTCTTAAGACTGCAACCGTTTTTTTGAACCAATGTTCCGATAG 677
QY 672 TGGACGATGGGCTTTCTTCCGAGGATGTTCTTAAGCCGAAAGTGGCGGTCATTGGCGC 731
DB 678 TGGACGATGGGCTTTCTTCCGAGGATGTTCTTAAGCCGAAAGTGGCGGTCATTGGCGC 737
QY 732 TGGCATTTCCGAGCTCGTGTGCAAAAGAACTGCTTCAATGCTGGGTTAGACGATGTTAC 791
DB 738 TGGCATTTCCGAGCTCGTGTGCAAAAGAACTGCTTCAATGCTGGGTTAGACGATGTTAC 797
QY 792 AATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGGTCAATGCTTTCCAGGACGC 851
DB 798 AATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGGTCAATGCTTTCCAGGACGC 857
QY 852 TCCAGTGTGTGGCGCAAAATGGGGGGATGCGATTTCTCTGCTGCTGATTTGCTGCTTT 911
DB 858 TCCAGTGTGTGGCGCAAAATGGGGGGATGCGATTTCTCTGCTGCTGATTTGCTGCTTT 917
QY 912 TTTCTTCTCGAGGCTTACGCGCTGTCTTGCATGAGGCGGTTCCCAATCCCGCACAGT 971
DB 918 TTTCTTCTCGAGGCTTACGCGCGTGTCTTGCATGAGGCGGTTCCCAATCCCGCACAGT 977
QY 972 CGACACTTAATCTGCTTACCAAGGCTTCAATATGATGGAAGCGGGGAGCTGCGAC 1031
DB 978 CGACACTTAATCTGCTTACCAAGGCTTCAATATGATGGAAGCGGGGAGCTGCGAC 1037
  
```

```

QY 1032 GAAGCTGTTCCATCGCGTTTACACCGTTGGCGTGGCTTTCTTGAAGACGGTTTTCATGA 1091
DB 1038 GAAGCTGTTCCATCGCGTTTACACCGTTGGCGTGGCTTTCTTGAAGACGGTTTTCATGA 1097
QY 1092 GCGAGATTAATGTTGGTGGCTTCCGCTTCCGCTATTAATCAAGGCTTTGAATCAGACACAT 1151
DB 1098 GCGAGATTAATGTTGGTGGCTTCCGCTTCCGCTATTAATCAAGGCTTTGAATCAGACACAT 1157
QY 1152 TAGTGGGCTCANTGATCCCGGCAAAATTTGGCTGAACCGTTTGGGAGGGAGTCTTCTC 1211
DB 1158 TAGTGGGCTCANTGATCCCGGCAAAATTTGGCTGAACCGTTTGGGAGGGAGTCTTCTC 1217
QY 1212 TTCAGGATTAAGAGGATCTTCTTGGGCAACATCTCTGCTGTGTAACATGAGATTTT 1271
DB 1218 TTCAGGATTAAGAGGATCTTCTTGGGCAACATCTCTGCTGTGTAACATGAGATTTT 1277
QY 1272 CCTCATGATTTGGGACCTATTCAAGCTTAATGGAATAGATCTGGCGGGTTTGGTCCAGT 1331
DB 1278 CCTCATGATTTGGGACCTATTCAAGCTTAATGGAATAGATCTGGCGGGTTTGGTCCAGT 1337
QY 1332 TTTTGAAGGGGTTTATGAGATCTCTCGCTTGTGATCAACGATATGAAGAAATCA 1391
DB 1338 TTTTGAAGGGGTTTATGAGATCTCTCGCTTGTGATCAACGATATGAAGAAATCA 1397
QY 1392 GCGGATGTGCCCTGAAGAAATCTCAGAACTTCCACGTCGGATCGCATCTGAAAGTGTAA 1451
DB 1398 GCGGATGTGCCCTGAAGAAATCTCAGAACTTCCACGTCGGATCGCATCTGAAAGTGTAA 1457
QY 1452 CGGTGTGCTGTGAGCCAGCGCATATGCCATGTTCAGTCAAGTCAAGGCGCATTCAGAAAGAAA 1511
DB 1458 CGGTGTGCTGTGAGCCAGCGCATATGCCATGTTCAGTCAAGTCAAGGCGCATTCAGAAAGAAA 1517
QY 1512 GACAAAATTAAGATTAAGGCTTAAGACGGGATATCGAATTTATGATTAAGTGGTGT 1571
DB 1518 GACAAAATTAAGATTAAGGCTTAAGACGGGATATCGAATTTATGATTAAGTGGTGT 1577
QY 1572 CACATCTGACCTCGCAATATCCAACTCAGGCACTTCCGTCAGATGCGATACCAATATTTT 1631
DB 1578 CACATCTGACCTCGCAATATCCAACTCAGGCACTTCCGTCAGATGCGATACCAATATTTT 1637
QY 1632 TCAGGACCAAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATTCCTCAAACTCTT 1691
DB 1638 TCAGGACCAAGTGAACCAAGCGGTTATTAACAGCCATATGACAGATTCCTCAAACTCTT 1697
QY 1692 CCTGATGACTGAACGAAATTTCTGTTAGACATATCTCCGCTTGTGTCTCATGGA 1751
DB 1698 CCTGATGACTGAACGAAATTTCTGTTAGACATATCTCCGCTTGTGTCTCATGGA 1757
QY 1752 CGGATCGCAAAAGCAGTATTCCTGACTATGAGTGCAGAGATCC 1799
DB 1758 CGGATCGCAAAAGCAGTATTCCTGACTATGAGTGCAGAGATCC 1805
  
```

RESULT 9  
 AAN60970  
 ID AAN60970 standard; DNA; 2126 BP.  
 AC AAN60970;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 28-OCT-1991 (first entry)  
 XX  
 XX  
 DE TL-DNA region of A. rhizogene agropiline-type plasmid pRiHRI.  
 KW  
 XX Ri plasmid; plant plasmid.  
 OS  
 XX Agrobacterium rhizogenes HRI and A4.  
 PN EP204590-A.  
 XX  
 PD 10-DEC-1986.  
 XX

PF 21-APR-1986; 86EP-0400855.  
 XX 22-APR-1985; 85US-0725368.  
 XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
 PA Slightom JL, Telfer DA;  
 XX WPI, 1986-327378/50.  
 XX  
 PT Plant cell genetic modification - using pri T-DNA promoter to  
 PT control expression of heterologous foreign structure genes, e.g.  
 PT for phaseolin.  
 PS Disclosure; Fig 2; 86pp; English.  
 XX The TL-DNA sequence is that of an Agrobacterium Ri plasmid, it  
 CC contains sixteen open reading frames, bounded by eukaryotic  
 CC promoters, ribosome binding and polyadenylation site. The plasmid  
 CC sequence may be used for genetic manipulation of plants, conveying  
 CC phenotypes such as improved resistance to environmental stress, disease  
 CC and parasitism.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 SQ Sequence 21126 BP; 5359 A; 4901 C; 5308 G; 5558 T; 0 other;  
 Query Match 6.0%; Score 136.4; DB 7; Length 21126;  
 Best Local Similarity 45.9%; Pred. No. 1.8e-32;  
 Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;

QY 623 TCGACTGCTCTACGACTGAGACCGTTTTTTCACATGTCGATAGTGAAGATCG 682  
 DB 7273 TTGACCTGGGGTACGACTATCAGGTTTCTCCAGAGTGCATGCCATGATGCTCG 7332  
 QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGTCATTGGCGTGCATTTCCG 742  
 DB 7333 GGCATTATCCGGAAGCGGCACCACTAGTGTTCATGTAGAGAGGCGGCTGTCTG 7392  
 QY 743 GACTCGGTGGGCAAGAACTGCTTATGCTGGGGTGAAGCATGTTACATATAGAA 802  
 DB 7393 GCTCTGTGGTGCACAGAACTTGGCGCTGCCTGCGTCAAGAAATCACTTTTCGATA 7452  
 QY 803 CAAATGATCGTGTGAGGCA---GCTTGTGCATATGCTTCAGGAGCGCTCTAGT 859  
 DB 7453 CCGTGTATGATCCGATGTTTGGGCAATCGCGATCCAAAGCGGAGCTCACAGG 7512  
 QY 860 TCGTGGCCGAATGAGGCGCATGAGATTCTCTGCTGCATTTCTGTTTCTTCC 919  
 DB 7513 CTTTGAAGTGTGCTGCTGATGCTTCTCCGCAACCAACTTGTCTGTCACTATC 7572  
 QY 920 TCGAGCTTACGGCTGTCTTTCATGAGCGCTTCCCAATCCCGGACAGTGCACATT 979  
 DB 7573 TGAATTAAGTTTAATTCCTGACGCTTGTGCTTGTGCGGAGACACACACAG 7632  
 QY 980 ACTTGTCTACCAAGGCGCTCAATATACATGGAAGCGGAGCGTCCACCGAAGCTGT 1039  
 DB 7633 CACTATATTTCCCGCAAAACGCTACGATGACGCGGAGCAAGCTCCCGGAGTAT 7692  
 QY 1040 TCCATGCGCTTTTACAAAGCTTGGCGTGCCTTCTTGAAGACGCTTTTCATAGCGAGATA 1099  
 DB 7693 TTACAGGCGGTACATGCTGATGGAAGACACTACTCTACCAAGGTGTGAAGGAAATGSCA 7752  
 QY 1100 TTGTGTTGGCTTGGCTGTGCTATTAAGTCAAGGCTTGAATACAGACATTAAGTGGG 1159  
 DB 7753 GAGAGCTGATGCTCCGATGATATCTTTCATATGTTGAAGAGCGTCTGATGATAG 7812  
 QY 1160 CTCATGCTCTGCGCAAAATTTGGCTGAACGTTTCGGGAGGAGTCTTCTTCAAGGA 1219  
 DB 7813 CTCACAGAGACGCGACCTTGGCTCGAGAGTTCCGAAATTCATCTTTCATGCGGTTT 7872  
 QY 1220 TAGAGAGATCTTCTTGGGACACATCTCTGCTGTGT---AAACATGAGTTTCTTC 1275  
 DB 7873 TGGTGCAGATCTTCAAGCTGTGTAATTCAGAGTCTGCTGTGCAAGGATGCAAAACACCCC 7932

QY 1276 ATGATTGGACCTTATCAAGCTAATGGAATAGATCTGGCGGGGTTTGGTCCAGTTT 1335  
 DB 7933 ATGATTTCAAGGCTTTCCGGATCTAGAGTTGGGAATACGGCCGAGTTTGTCC-TATTAC 7991  
 QY 1336 GAAAGCGGTTTATTTAGATCTCTCCGCTTGTCTATCAACGATATGAAGAAATACAGCG 1395  
 DB 7992 AACGTGTTGTTTCAACGATCTGACTGATATATCAATGCTACAGAGGAGCCAGCAT 8051  
 QY 1396 ATGTGCTTGAAGAAATCTCAGAACTTCCACGTCGATGCGATCTGAAGTGTAAACGT 1455  
 DB 8052 CTTTCTATTGTGGGGTTCAACTTTTTCAGGCTCTGATCGCAATTAATATTCAGAAA 8111  
 QY 1456 GTGTCTGTGAGCCAGGCAATATGCTATGTTCAAGTCCAGGCGCATTTGAAGAAAGACA 1515  
 DB 8112 AGCCATGCGAAGACACGACTCTGTTTATCCGTCGTGGAATACCAAGGAGGCGGG 8171  
 QY 1516 AAAATTAAGATTAAGCTTAAAGCGGATATCTGAATTTATGATTAAGTGTGTACACA 1575  
 DB 8172 AGATTGAAGTATGCTTGAACACGCTCATTTGCGCTTTTGAACAGGTCATCATTTGGC 8231  
 QY 1576 TCTGACTGCGCAATATCCAACTCAGGCAATTGCTGACATGCGATACCAATATTTTCAG 1635  
 DB 8232 GGCAGTGTGAGGCGCTACAGTTGATTAACAGACTGCGGAGATGAGACTTCTTCAGC 8291  
 QY 1636 GCACCACTGAACCAAGCGGTGATTAACAGCATATGACAGATCCGCAAACTCTTCCTG 1695  
 DB 8292 TACATATATGAACCCCGCTCGAATCTGCTGCTGCTGCTCAATTCAGACATCTTCATG 8351  
 QY 1696 ATGACTGAACGAATAATCTGTGTGAACCATATCTCTCCGCTTGTGCTCTCATGACGG 1755  
 DB 8352 GTACAGCAACGAATAAGTTTGGGTAACTCCGGATCCAGCATGATATAGAACCATGAG 8411  
 QY 1756 ATGCAAAAGCACTGTATTTGCTGACTATGAGTGCAGAGATCCGAATGTAAGTCTTA 1815  
 DB 8412 CTTTGTCCGAGTGTGTGCAATTCATGCAATCGAATCG-----CCAGCTGAGAGGCGCTT 8465  
 QY 1816 GTGCTCATGATTAATCATGAGGAGGAGCGACTCCCAAGCTGTGGCGGTCCCGGACAA 1875  
 DB 8466 GTGCTTTCACATGATGTTTGAATGATCTATATCCCGGCGCATGACATGACAAAG 8525  
 QY 1876 AAAGAGCGATTAATGTCTGCTGCGGAGCGCAATTTTCAGATCTTTCCGCGCTTGGCCAG 1935  
 DB 8526 AAGGAGCGGTCTTGAATTTGTCAGGAGCTTGTCTGCTCTTCTTCAACTGCGCTGT 8585  
 QY 1936 CACCTATTTCTGCTGCGCTGATTAACACCAAAATGT 1973  
 DB 8586 CACCTGTCTCCAGTCAACGAAGCTACGAACGATATGT 8623

RESULT 10  
 ABL13666  
 ID ABL13666 standard; cDNA; 4734 BP.  
 XX  
 AC ABL13666;  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35480.  
 XX  
 KW Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 OS Drosophila melanogaster.  
 XX  
 PN MO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX  
 PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX P-PSDB; ABB69563.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX  
XX Claim 1; SEQ ID NO 35480; 21bp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
XX sequences (AB101840-AB16175) and the encoded proteins  
XX (ABB57737-ABB72072).  
XX The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 4734 BP; 1394 A; 978 C; 1048 G; 1314 T; 0 other;  
SQ  
Query Match 1.8%; Score 39.8; DB 23; Length 4734;  
Best Local Similarity 52.0%; Pred. No. 0.28;  
Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 690 TCCGGAGAGATTCTCTAAGCCGAAAGTGGCGGTATGGCGCTGCTTCCGACTCGT 749  
DB 1219 TTCAAGGAATATATACCAACGTTAATCGATCGATGAGCTGGAATGGCTGCTATC 1278  
QY 750 GGTGGCAAGCAAGTCTTCAATGCTGGGATGAGCATGTTACATATATGAAAGCAAGTGA 809  
DB 1279 GGCTGCAATACCTTTTGGAAAAGCGCTGTGACGATTTCTTATTTGAGAGCAGGGG 1338  
QY 810 TCGTGTGGAGGCAAGCTTTGTCACATGCTTTGAGGAGCGCTCTAGTGT 860  
DB 1339 CCGGGTGGGGGAGCAGCATTTGTGTCATACCGCTAGCAATATCAGAAAGT 1389  
RESULT 11  
AB15662  
ID ABL15662 standard; cDNA; 4800 BP.  
XX  
AC ABL15662;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 41468.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; ss.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
XX  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX

DR WPI; 2001-656860/75.  
DR P-PSDB; ABB71559.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX  
XX Claim 1; SEQ ID NO 41468; 21bp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
XX sequences (AB101840-AB16175) and the encoded proteins  
XX (ABB57737-ABB72072).  
XX The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 4800 BP; 1265 A; 1116 C; 1141 G; 1278 T; 0 other;  
SQ  
Query Match 1.8%; Score 39.8; DB 23; Length 4800;  
Best Local Similarity 52.0%; Pred. No. 0.28;  
Matches 89; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 690 TCCGGAGAGATTCTCTAAGCCGAAAGTGGCGGTATGGCGCTGCTTCCGACTCGT 749  
DB 4496 TTCAAGGAATATATACCAACGTTAATCGATCGATGAGCTGGAATGGCTGCTATC 4555  
QY 750 GGTGGCAAGCAAGTCTTCAATGCTGGGATGAGCATGTTACATATATGAAAGCAAGTGA 809  
DB 4556 GGCTGCAATACCTTTTGGAAAAGCGCTGTGACGATTTCTTATTTGAGAGCAGGGG 4615  
QY 810 TCGTGTGGAGGCAAGCTTTGTCACATGCTTTGAGGAGCGCTCTAGTGT 860  
DB 4616 CCGGGTGGGGGAGCAGCATTTGTGTCATACCGCTAGCAATATCAGAAAGT 4666  
RESULT 12  
AB13667  
ID ABL13667 standard; cDNA; 1872 BP.  
XX  
AC ABL13667;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 35483.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; ss.  
XX  
XX Drosophila melanogaster.  
XX  
XX WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
XX  
XX 11-JUL-2000; 2000US-0614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
XX  
XX P-PSDB; ABB69564.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -  
XX  
PS Claim 1; SEQ ID NO 35483; 21bp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABU16176-ABU16175), expressed DNA  
CC sequences (ABB5773-ABB572072).  
CC (ABB5773-ABB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1872 BP; 499 A; 477 C; 524 G; 372 T; 0 other;  
XX  
Query Match 1.7%; Score 39.2; DB 23; Length 1872;  
Best Local Similarity 52.4%; Pred. No. 0.25; Mismatches 0; Gaps 0;  
Matches 86; Conservative 0; Indels 0; Gaps 0;  
XX  
QY 690 TCCGAGAGATGTTCTTAAGCCGAAAGTGCGCGTTCATTCGCGATTTCCGACTCGT 749  
DB 219 TTCAGGGAATTAATCAACGTTAGATCGATCATTCGAGCTGGAATGGCTGCTATC 278  
XX  
QY 750 GGTGGCAACGAACTGCTTCATGCTGGGGTAGAGATGATTAATATAGCAAGTGA 809  
DB 279 GGCTGCAAAATCACCTTTTGCAAAACGCTGTGAGATTTCTTAATTCGAGGACGCGG 338  
XX  
QY 810 TCGGTGTGAGAGCAAGCTTTGGTCAATGCTTTCAGGAGCGCTC 853  
DB 339 CCGGGTGGCGGAGCGCATTTGTTCATACCGCTCAGCAATATATC 382  
XX  
RESULT 13  
AAK15122  
ID AAK15122 standard; cDNA to mRNA; 1575 BP.  
XX  
AC AAK15122;  
XX  
DT 23-APR-1999 (first entry)  
XX  
DE Nucleic acid encoding an apoptosis inducing protein.  
XX  
KW Cell death; apoptosis; inhibition; proliferation; cancer cell;  
KW apoptosis inducing protein; AIP; chub mackerel; anticancer; ss.  
XX  
OS *Scomber japonicus*.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..1575  
XX /\*tag= a  
XX  
PN WO9852972-A1.  
XX  
PD 26-NOV-1998.  
XX  
PF 22-MAY-1998; 98WO-JP02261.  
XX  
PR 23-MAY-1997; 97JP-0133549.  
XX  
PA (TENS-) TENSBI SUISAN CO LTD.  
XX  
PI Iwamoto M, Jung S;  
XX  
XX WPI; 1999-070139/06.  
XX  
XX P-PDB; AAM96805.  
XX  
PT Apoptosis inducing protein from mackerel and gene encoding it - for  
XX use as anticancer agents  
XX  
PS Claim 5; Page 32-34; 47pp; Japanese.

XX  
CC The present sequence encodes a protein which induces cell death  
CC (apoptosis) and inhibits the proliferation of cancer cells. The  
CC protein (apoptosis inducing protein, AIP) is isolated from chub  
CC mackerel. The protein can be used as an anticancer agent and as a  
CC reagent for study of the mechanisms of apoptosis in vitro.  
XX  
SQ Sequence 1575 BP; 436 A; 358 C; 406 G; 375 T; 0 other;  
XX  
Query Match 1.7%; Score 38.6; DB 20; Length 1575;  
Best Local Similarity 66.1%; Pred. No. 0.35;  
Matches 72; Conservative 0; Mismatches 34; Indels 3; Gaps 1;  
XX  
QY 713 AACTGGCGGTCATTCGCGCTGCGATTTCCGAGTCTGTCGCAACGACTGCTCATG 772  
DB 182 ATGTGTTTATAGTCGAGCTGCGATGCGGACTGACGCGGCAATTAATGCAAGACG 241  
XX  
QY 773 CTGGGCTAGACGATGTTACATATATATGAAGCAAGTATGTTGAGG 821  
DB 242 CAGG---ACACACGATACCATTTGAGGCTATATCATCTGTTGAGG 287  
XX  
RESULT 14  
ABZ56743  
ID ABZ56743 standard; cDNA; 468 BP.  
XX  
AC ABZ56743;  
XX  
DT 28-MAR-2003 (first entry)  
XX  
DE *Aspergillus oryzae* polynucleotide SEQ ID NO 5856.  
XX  
KW *Aspergillus oryzae*; fermentation; fungus; industrial; EST;  
KW expressed sequence tag; gene; ss.  
XX  
OS *Aspergillus oryzae*.  
XX  
PN WO200279476-A1.  
XX  
PD 10-OCT-2002.  
XX  
PF 22-MAR-2002; 2002WO-IB00890.  
XX  
PR 30-MAR-2001; 2001JP-0098371.  
XX  
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
PA (NARE-) NAT RES INST BREWING.  
PA (NORQ) NAT FOOD RES INST MIN AGRIC.  
XX  
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;  
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;  
XX  
XX WPI; 2003-046817/04.  
XX  
PT Detection of expression of specific *Aspergillus* genes for monitoring  
PT the fermentation and growth conditions of the fungus, using DNA probes  
XX  
PS Claim 1; SEQ ID NO 5856; 48pp + Sequence Listing; Japanese.  
XX  
CC The invention relates to a polynucleotide having any of 6006 specific  
CC sequences (ABZ50888-ABZ58893), which are expressed by a fungus under  
CC specific culture conditions including one or more of eutrophic,  
CC oligotrophic, solid, early germination, alkaline, high temperature, low  
CC temperature or maltose culture or polynucleotides stringently hybridizing  
CC to these sequences. The polynucleotides are useful for monitoring the  
CC progress of fermentation and the growth conditions of a fungus,  
CC especially of *Aspergillus oryzae* which is widely used in industrial  
CC fermentation. Also monitoring for fungal contamination.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 468 BP; 124 A; 141 C; 111 G; 90 T; 2 other;

Query Match 1.7%; Score 38.4; DB 25; Length 468;  
Best Local Similarity 53.8%; Pred. No. 0.19;  
Matches 100; Conservative 1; Mismatches 82; Indels 3; Gaps 1;

QY 706 AACCCGAAATGCGCGCTTATGCGCTGCGATTTCCGCACTCGTGTGCAAGCAACTG 765  
DB 164 AACACCAATATCCGATCATCGCGCGCGCATCTCAGGCTTCAAGCCGCTCTCTC 223  
QY 766 CTTCATGCTGGGGTGAAGCATGTATACATATATGAAGCAAGTATCTGTGAGGCAAG 825  
DB 224 GATAGCGCTCGGAATCCCAATGGAATACTCGAAGCGAGTAAACGGGTAGGAGGCGGT 283  
QY 826 CTTCGTCACATGCTTTCAGGAGCGCTCTAGTGTCTGTCGCGCAATGCGGCGATCCGA 885  
DB 284 TTCGGAACCAATCTGTAGGGAACACACG--GGCAATGGGCGGARAATGGGGCCATGCGG 340  
QY 886 TTTCTCT 891  
DB 341 CTGCTCT 346

## RESULT 15

AAC05199  
ID AAC05199 standard; cDNA; 273 BP.

AC AAC05199;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 9274.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 9274; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.

SQ Sequence 273 BP; 53 A; 73 C; 66 G; 80 T; 1 other;

Query Match 1.7%; Score 37.8; DB 21; Length 273;  
Best Local Similarity 50.8%; Pred. No. 0.21;  
Matches 90; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 867 CGAATGAGGAGGAGATCGATTTCTCTCTGCTGCTTCTGTTCTTCTCTGAGCG 926  
DB 97 CGATCTGGGCTCGATGCAACCTCTGCTCCAGGTTCAAGCAATTTCTGCTGGGAG 156  
QY 927 TTACGCTCTGCTTCTGATGAGCGGTTCCCAATCCGGCAGATGACACTTACTTGT 986  
DB 157 AGACGGGGTTTCAACATGTTGGCAAGCTGCTCGAACCCCTGACCTCAAGTATCAGC 216  
QY 987 CTACCAAGGCGTCCATATACATGTGGAAGCCGGGACGTCGACCGAGCTGTCA 1043  
DB 217 CCACCTGGCTTCCCAAGTGTGAGATTACAGGTGTGATCCACTGCACTCGGTCCA 273

Search completed: November 23, 2003, 21:01:38

Job time : 607.746 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 20:47:41 ; Search time 8367.46 Seconds  
(without alignments)  
11088.547 Million cell updates/sec

Title: US-09-434-837-10

Perfect score: 2268  
Sequence: 1 atgcacagcttcacactctctcct.....actacgcagtagaattag 2268

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_pac: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_da: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pac: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vi: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pia: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_man: \*  
37: em\_htg\_vtc: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID        | Description          |
|------------|--------|-------------|--------|--------------|----------------------|
| 1          | 2230.6 | 98.4        | 24595  | 1 ATACH5     | X00493 Agrobacteri   |
| 2          | 2230.6 | 98.4        | 24595  | 6 BD016312   | BD016312 Method of   |
| 3          | 2230.6 | 98.4        | 24595  | 6 E00404     | E00404 Ti plasmid    |
| 4          | 2230.6 | 98.4        | 24595  | 6 E00546     | E00546 DNA fragment  |
| 5          | 2230.6 | 98.4        | 194140 | 1 AF242881   | AF242881 Agrobacte   |
| 6          | 2225.8 | 98.1        | 2591   | 1 T1PMS2     | K0254 Ti plasmid     |
| 7          | 2044   | 90.1        | 9931   | 1 AE007927   | AE007927 Agrobacte   |
| 8          | 2044   | 90.1        | 10377  | 1 AE009419   | AE009419 Agrobacte   |
| 9          | 2044   | 90.1        | 206479 | 1 AB016260   | AB016260 Agrobacte   |
| 10         | 2042.4 | 89.6        | 4660   | 1 AF126446   | AF126446 Agrobacte   |
| 11         | 2032.8 | 89.6        | 4660   | 1 AB025110   | AB025110 Agrobacte   |
| 12         | 2016.2 | 88.9        | 29524  | 1 ATTU237588 | ATTU237588 Agrobacte |
| 13         | 1880.8 | 82.9        | 10200  | 1 ATTH1A8H   | X56185 A. tumefacie  |
| 14         | 1880.8 | 82.9        | 15463  | 1 AV083987   | U83987 Agrobacteri   |
| 15         | 1880.8 | 82.9        | 15463  | 1 AV083986   | U83986 Agrobacteri   |
| 16         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 17         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 18         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 19         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 20         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 21         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 22         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 23         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 24         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 25         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 26         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 27         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 28         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 29         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 30         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 31         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 32         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 33         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 34         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 35         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 36         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 37         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 38         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 39         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 40         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 41         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 42         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 43         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 44         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |
| 45         | 1880.8 | 82.9        | 14960  | 1 AF126447   | AF126447 Agrobacte   |

## ALIGNMENTS

RESULT 1  
LOCUS ATACH5 24595 bp DNA linear BCT 10-FEB-1999  
DEFINITION Agrobacterium tumefaciens Ti plasmid pTi15955-T-DNA region.  
ACCESSION X00493 J05108 X00282  
VERSION X00493.1 GI:39062  
KEYWORDS octopine synthetase; plasmid; synthetase; terminal repeat;  
unidentified reading frame.  
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)  
ORGANISM Agrobacterium tumefaciens  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
REFERENCE 1 (bases 1 to 24595)  
AUTHORS Barker R.F., Idler K.B., Thompson D.V. and Kemp J.D.  
TITLE Nucleotide sequence of the T-DNA region from the Agrobacterium

|  |  |
|--|--|
| JOURNAL<br>REFERENCE<br>AUTHORS<br>TITLE | tumefaciens octopine Ti plasmid pTi15955<br>2 (bases 602 to 14237)<br>Gjelen,U., De Beuckeleer,M., Seurinck,J., Deboeck,F., De Greve,H.,<br>Lemmers,M., Van Montagu,M. and Schell,J.<br>The complete nucleotide sequence of the Ti-DNA of the Agrobacterium<br>tumefaciens plasmid pTiAch5<br>EMBO J. 3 (4), 835-846 (1984)  |
| JOURNAL<br>MEDLINE<br>PUBMED             | 84207942<br>6327292<br>3 (bases 1 to 24595)<br>Turk,S.C., Nester,E.W. and Hooykaas,P.J.<br>The <i>vira</i> promoter is a host-range determinant in <i>Agrobacterium</i><br><i>tumefaciens</i><br>Mol. Microbiol. 7 (5), 719-724 (1993)   |
| JOURNAL<br>PUBMED                        | 93225814<br>8469115<br>4 (bases 1 to 24595)<br>Guevara-Garcia,A., Mosqueda-Cano,G., Arguello-Astorga,G.,<br>Simpson,J. and Herrera-Betrelle,L.<br>Tissue-specific and wound-inducible pattern of expression of the<br>manopine synthase promoter is determined by the interaction<br>between positive and negative cis-regulatory elements<br>Plant J. 4 (3), 495-505 (1993)   |
| JOURNAL<br>MEDLINE<br>PUBMED             | 94035196<br>8220492<br>Location/Qualifiers<br>1. .24595<br>/organism="Agrobacterium tumefaciens"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:358"<br>/plasmid="Ti"<br>909. .23782<br>/note="T-region"<br>909. .932<br>/note="direct repeat A"<br>1337<br>/note="C is T in [2]"<br>/citation=[2]<br>1661. .2179<br>/note="open reading frame 1"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25163.1"<br>/db_xref="GI:39063"<br>/db_xref="SWISS-PROT:P04028"<br>/translation="MYDGOPIFNIDSSNLODRRLKVLHTENAVYSSAORSILAS<br>ORWNINLINTDVPIDPADDEVKCSRKVACCPDPTDLPDILNVSLILYYVCSQENR<br>RVAQGFYDGVSDGNAVISTVPPRYAGITKQTRMLWQKKWONTSKETHDADYIALL<br>PNRFRKIQISHT" |
| CDS                                      | complement(1736..2113)<br>/note="open reading frame 2"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25164.1"<br>/db_xref="GI:39064"<br>/db_xref="SPRMBL:Q44387"<br>/translation="MCFPACILPNFPLPOPHSLFCYSGIGDGGDNCAVGNVAVS<br>LUCIPHPILKRAINVQORYIEYIKGYICVMAARNFARALDHFVFSWINGNISINDEI<br>YPRALSYETSLCRTAISILCVN"  |
| CDS                                      | complement(1736..2113)<br>/note="open reading frame 2"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25164.1"<br>/db_xref="GI:39064"<br>/db_xref="SPRMBL:Q44387"<br>/translation="MCFPACILPNFPLPOPHSLFCYSGIGDGGDNCAVGNVAVS<br>LUCIPHPILKRAINVQORYIEYIKGYICVMAARNFARALDHFVFSWINGNISINDEI<br>YPRALSYETSLCRTAISILCVN"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/note="T is CG in [2]"<br>/citation=[2]<br>complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"   |
| CDS                                      | complement(2907..3287)<br>/note="open reading frame 3"<br>/codon_start=1<br>/transl_table=11<br>/protein_id="CAA25165.1"<br>/db_xref="GI:39065"<br>/db_xref="SWISS-PROT:P03867"  |
| conflict                                 | 1948<br>/note="C is T in [2]"<br>/citation=[2]<br>2146<br>/  |

| Query Match | Best Local Similarity | Matches 2258;  | Conservative | 0; | Mismatches | 9; | Indels | 2; | Gaps | 2; |
|-------------|-----------------------|--|--------------|----|------------|----|--------|----|------|----|
| QY          | 1                     | ANGTCAGCTTACCCCTCCCTGATTAACAGAGCGCATCATCTCCCAACCAAAATGCTGAT  | 60           |    |            |    |        |    |      |    |
| Db          | 5809                  | ANGTCAGCTTACCCCTCCCTGATTAACAGAGCGCATCATCTCCCAACCAAAATGCTGAT  | 5868         |    |            |    |        |    |      |    |
| QY          | 61                    | CTGACAAATGTCGATTAAGCGCGATGAATTGACCGCAGAGGTTTCCGATGCTTCTTGAA  | 120          |    |            |    |        |    |      |    |
| Db          | 5869                  | CTGACAAATGTCGATTAAGCGCGATGAATTGACCGCAGAGGTTTCCGATGCTTCTTGAA  | 5928         |    |            |    |        |    |      |    |
| QY          | 121                   | CGAGAAGCTTAAAGGGAAGAGATTACTCAAAAGCTTCCACCGAGTGCAGCGCTGGTTA   | 180          |    |            |    |        |    |      |    |
| Db          | 5929                  | CGAGAAGCTTAAAGGGAAGAGATTACTCAAAAGCTTCCACCGAGTGCAGCGCTGGTTA   | 5988         |    |            |    |        |    |      |    |
| QY          | 181                   | GCTTGCAAAAGGCTGCGCGATGTCGCTTCCCGAGATCTCAAGCTGTGTGAAAAGTGTACA | 240          |    |            |    |        |    |      |    |
| Db          | 5989                  | GCTTGCAAAAGGCTGCGCGATGTCGCTTCCCGAGATCTCAAGCTGTGTGAAAAGTGTACA | 6048         |    |            |    |        |    |      |    |
| QY          | 241                   | GTTCTCTCCGCTTATATCTATATTGGCAAAAGAAATTTGGGGCGGATCTTGAATCGAAA  | 300          |    |            |    |        |    |      |    |
| Db          | 6049                  | GTTCTCTCCGCTTATATCTATATTGGCAAAAGAAATTTGGGGCGGATCTTGAATCGAAA  | 6108         |    |            |    |        |    |      |    |
| QY          | 301                   | CTTGGGCGGCGGCAACAGTAGTGTCTGCTTGGCATCGACTTGGACACCATTTGCATG    | 360          |    |            |    |        |    |      |    |
| Db          | 6109                  | CTTGGGCGGCGGCAACAGTAGTGTCTGCTTGGCATCGACTTGGACACCATTTGCATG    | 6168         |    |            |    |        |    |      |    |
| QY          | 361                   | GATTTCTCCGAAGCAACAATCAAGACCTGTTTTGTTGCTGAGCGGTAAAGATGTGCA    | 420          |    |            |    |        |    |      |    |
| Db          | 6169                  | GATTTCTCCGAAGCAACAATCAAGACCTGTTTTGTTGCTGAGCGGTAAAGATGTGCA    | 6228         |    |            |    |        |    |      |    |

|    |      |   |      |
|----|------|---|------|
| QY | 421  | CCGATATGATCTTAGCATTTGTGGCCATTTCAATCTCTAAGACGCGGGCTTTCGAAC       | 480  |
| Db | 6229 | CCGATATGATCTTAGCATTTGTGGCCATTTCAATCTCTAAGACGCGGGCTTTCGAAC       | 6288 |
| QY | 481  | CTGCCAATGCCGCTGTACGAGAATGGCAGATGAAATGCGTTACCGGGTTTACATAAC       | 540  |
| Db | 6289 | CTGCCAATGCCGCTGTACGAGAATGGCAGATGAAATGCGTTACCGGGTTTACATAAC       | 6348 |
| QY | 541  | CTTGAAGGGGCGGTGCCATTTTGACATGGTACTTATGGTCCGAAACCTGATCTGAAGGGT    | 600  |
| Db | 6349 | CTTGAAGGGGCGGTGCCATTTTGACATGGTACTTATGGTCCGAAACCTGATCTGAAGGGT    | 6408 |
| QY | 601  | TCGGCAGAGTTTCCTTTCCAAACAATCGACTTGTGCTACGACTCGACGCGTTTTTTTGAACCA | 660  |
| Db | 6409 | TCGGCAGAGTTTCCTTTCCAAACAATCGACTTGTGCTACGACTCGACGCGTTTTTTTGAACCA | 6466 |
| QY | 661  | TGTTCCGATAGTGAACGAGTCGGCTTCTTTCGGAGAGTGTTCCTAGCCGAAAGTGGC       | 720  |
| Db | 6469 | TGTTCCGATAGTGAACGAGTCGGCTTCTTTCGGAGAGTGTTCCTAGCCGAAAGTGGC       | 6522 |
| QY | 721  | GTCATTTGGCGGTGGCATTTCCGACCTCGTGGTGGCAAAAGAACTGCTCATGCTGGGGTA    | 780  |
| Db | 6529 | GTCATTTGGCGGTGGCATTTCCGACCTCGTGGTGGCAAAAGAACTGCTCATGCTGGGGTA    | 6588 |
| QY | 781  | GACGATGTTACAATATATGAAAGCAAGTATCGTGTGGAGCAAGCTTTGGTTCACATCT      | 840  |
| Db | 6589 | GACGATGTTACAATATATGAAAGCAAGTATCGTGTGGAGCAAGCTTTGGTTCACATCT      | 6646 |
| QY | 841  | TTTCAGGGACGCTCTCTAGTGTCTGTGGCCGAAATGGGGGCGATGCGATTTCTCTCTGCTCA  | 900  |
| Db | 6649 | TTTCAGGGACGCTCTCTAGTGTCTGTGGCCGAAATGGGGGCGATGCGATTTCTCTCTGCTCA  | 6708 |
| QY | 901  | TTTCGCTGTTTTCCTTCTCTCGACGCTGATCGGCGTCTTGATGAGGCGCTTCCCAAT       | 960  |
| Db | 6709 | TTTCGCTGTTTTCCTTCTCTCGACGCTGATCGGCGTCTTGATGAGGCGCTTCCCAAT       | 6766 |
| QY | 961  | CCCGGCACAGTCGACACTTACTTGGCTCTCAACGAGCGTCCAAATCATGTGAAAGCCGGG    | 1020 |
| Db | 6769 | CCCGGCACAGTCGACACTTACTTGGCTCTCAACGAGCGTCCAAATCATGTGAAAGCCGGG    | 6828 |
| QY | 1021 | CAGCTGCACACGGAAGCTGTTCCATCGCGGTTTACAACGGTTTGGCGGCTTCTTGAAGAC    | 1080 |
| Db | 6829 | CAGCTGCACACGGAAGCTGTTCCATCGCGGTTTACAACGGTTTGGCGGCTTCTTGAAGAC    | 6888 |
| QY | 1081 | GGTTTTCATGAGCGAGATATTTGTGTGGCTTCCGCTGTCCGTAATCTCAGAGCCTTGAA     | 1140 |
| Db | 6889 | GGTTTTCATGAGCGAGATATTTGTGTGGCTTCCGCTGTCCGTAATCTCAGAGCCTTGAA     | 6948 |
| QY | 1141 | TCAGGACACATTTAGTGGGCTCATGACTCTCTGGCAAAATTTGGCTGAAACCGTTTGGGAGG  | 1200 |
| Db | 6949 | TCAGGAGACATTTAGTGGGCTCATGACTCTCTGGCAAAATTTGGCTGAAACCGTTTGGGAGG  | 7008 |
| QY | 1201 | GAGTCTTCTCTTTCAGGGATGAGAGAGTCTTTCTGGGACACATTCCTCTGATGTGT-AA     | 1258 |
| Db | 7009 | GAGTCTTCTCTTTCAGGGATGAGAGAGTCTTTCTGGGACACATTCCTCTGATGTGTAAA     | 7066 |
| QY | 1260 | ACATGAGATTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAAATGAGATCTTGGCGGG   | 1315 |
| Db | 7069 | ACATGAGATTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAAATGAGATCTTGGC-GG   | 7122 |
| QY | 1320 | GTTTGTGTCAGTTTTCGAAAGCGGGTTTATGGAATCTCCGCTTGGCTCATCAACGAGTA     | 1379 |
| Db | 7128 | GTTTGTGTCAGTTTTCGAAAGCGGGTTTATGGAATCTCCGCTTGGCTCATCAACGAGTA     | 7188 |
| QY | 1380 | TGAAGAAATATCAGCGGATGTGACCCTCGAAGGAATCTCAGAACTTCCACGTCGGATCGCATC | 1439 |
| Db | 7188 | TGAAGAAATATCAGCGGATGTGACCCTCGAAGGAATCTCAGAACTTCCACGTCGGATCGCATC | 7244 |
| QY | 1440 | TGAAGTGTATACGGTGTGTCTGTGAGCCAGCGCATATGCGATGTTCAGATCAGAGCGAT     | 1499 |
| Db | 7248 | TGAAGTGTATACGGTGTGTCTGTGAGCCAGCGCATATGCGATGTTCAGATCAGAGCGAT     | 7307 |

QY 1500 TCAGAGGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATCTTATGA 1559  
 DB 7308 TCAGAGGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATCTTATGA 7367  
 QY 1560 TAAAGTGTGTGTCATCTGAGCTCGCAATATCAACTCAGGAGTTGCTGACATGCGA 1619  
 DB 7368 TAAAGTGTGTGTCATCTGAGCTCGCAATATCAACTCAGGAGTTGCTGACATGCGA 7427  
 QY 1620 TACCAATATTTTAAAGCAACAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 1679  
 DB 7428 TACCAATATTTTAAAGCAACAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 7487  
 QY 1680 GTCAAACTCTTCTGTAGTGAAGCAAAATTTGTTAGAACCATATCCCGCTTGG 1739  
 DB 7488 GTCAAACTCTTCTGTAGTGAAGCAAAATTTGTTAGAACCATATCCCGCTTGG 7547  
 QY 1740 TGTCTCATGAGCGGATTCGCAAAAGCATGTATTTGCTGAGCTATGATGACAGATCC 1799  
 DB 7548 TGTCTCATGAGCGGATTCGCAAAAGCATGTATTTGCTGAGCTATGATGACAGATCC 7607  
 QY 1800 GAATGTGAAGGTCTAGTGTCTCATGATTAATCAATGAGAGACACTCCCAAGCTGTT 1859  
 DB 7608 GAATGTGAAGGTCTAGTGTCTCATGATTAATCAATGAGAGACACTCCCAAGCTGTT 7667  
 QY 1860 GGGCGTCCCGACAAAGAGCATTAATGCTGCGGAGCGCAATTTGAGATCTTT 1919  
 DB 7668 GGGCGTCCCGACAAAGAGCATTAATGCTGCGGAGCGCAATTTGAGATCTTT 7727  
 QY 1920 CCGGCGTTTGGCCGAGCACTATTTCTGCTGCGCTGATTAACGCAAAATGTTATCA 1979  
 DB 7728 CCGGCGTTTGGCCGAGCACTATTTCTGCTGCGCTGATTAACGCAAAATGTTATCA 7787  
 QY 1980 ACATGATGTGCTTCAACAGCAATTCGCGGAGGAGCTTTCAAACTCAACCGGCTGTGA 2039  
 DB 7788 ACATGATGTGCTTCAACAGCAATTCGCGGAGGAGCTTTCAAACTCAACCGGCTGTGA 7847  
 QY 2040 GGATTTTATTTCTGAAGAACTTTCTTCAAGCACTGAGCACTGATTAACCGGAGT 2099  
 DB 7848 GGATTTTATTTCTGAAGAACTTTCTTCAAGCACTGAGCACTGATTAACCGGAGT 7907  
 QY 2100 TTACTTGGCGGTTGCAATGTTCTTCAACAGTGAATGAGTGGAGGTTGATTTCAAC 2159  
 DB 7908 TTACTTGGCGGTTGCAATGTTCTTCAACAGTGAATGAGTGGAGGTTGATTTCAAC 7967  
 QY 2160 CGGCTGTAAGCGGCTCTGTGCAATTAATCCAAATTTGAGAGCACTTTTGGCAAGGCGCA 2219  
 DB 7968 CGGCTGTAAGCGGCTCTGTGCAATTAATCCAAATTTGAGAGCACTTTTGGCAAGGCGCA 8027  
 QY 2220 TCCCTCGAACACTCTTGAAGAGATTAATTAACCGAGTAGAATTTAG 2268  
 DB 8028 TCCCTCGAACACTCTTGAAGAGATTAATTAACCGAGTAGAATTTAG 8076  
 RESULT 2  
 BD016312 24595 bp DNA linear PAT 27-AUG-2002  
 LOCUS Method of promoting plant transcription by using octopine T-DNA  
 DEFINITION  
 PROMOTER.  
 BD016312  
 ACCESSION BD016312 GI:22557450  
 VERSION UP 2001190289-A/1.  
 KEYWORDS Agrobacterium tumefaciens (Rhizobium radiobacter)  
 SOURCE Agrobacterium tumefaciens  
 ORGANISM Agrobacterium tumefaciens (Rhizobium radiobacter)  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 REFERENCE 1 (bases 1 to 24595)  
 AUTHORS Barker, R. F. and Kemp, J. D.  
 TITLE Method of promoting plant transcription by using octopine T-DNA  
 JOURNAL MYCOGEN PLANT SCIENCE INC  
 COMMENT OS Agrobacterium tumefaciens  
 PN JP 2001190289-A/1  
 PD 17-JUL-2001

PF 22-NOV-2000 JP 2000356816  
 PR 18-NOV-1983 US 553786  
 PI RICHARD F BARKER, JOHN D KEMP  
 PC C12N15/09, A01H5/00, C12N15/10, C12N15/00, C12N5/00 CC Method of  
 promoting plant transcription by using octopine T-CC  
 DNA promoter  
 FH Key  
 FT source 1. 24595  
 FT location/Qualifiers  
 location/Qualifiers  
 1. 24595  
 /organism="Agrobacterium tumefaciens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:358"  
 BASE COUNT 6534 a 5509 c 5793 g 6759 t  
 ORIGIN  
 Query Match 98.4%; Score 2230.6; DB 6; Length 24595;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;  
 QY 1 ATGTCAAGCTTCACTCTCTTGTATTAACAGTGCATCATCTCCCAACCAAAATGTTGAT 60  
 DB 5809 ATGTCAAGCTTCACTCTCTTGTATTAACAGTGCATCATCTCCCAACCAAAATGTTGAT 5868  
 QY 61 CTGCAATGTTGATTAAGGCGGATGTAATTGACCGAGGTTTCCGATGCTTTAGAA 120  
 DB 5869 CTGCAATGTTGATTAAGGCGGATGTAATTGACCGAGGTTTCCGATGCTTTAGAA 5928  
 QY 121 CGAAGAGCTTCTAGGGAAGAGATTACTCAAGTCCACCGAGTGCAGCGTGGTTA 180  
 DB 5929 CGAAGAGCTTCTAGGGAAGAGATTACTCAAGTCCACCGAGTGCAGCGTGGTTA 5988  
 QY 181 GCTTGCAAAAGCTGCGCATGTGCTTCCCGAATCTCACTGTGGAAGTGA 240  
 DB 5989 GCTTGCAAAAGCTGCGCATGTGCTTCCCGAATCTCACTGTGGAAGTGA 6048  
 QY 241 GTTCTCTCGCTTATTAATTAATTTGGCAAGAAATCTGGGGGGGATTAATGAAA 300  
 DB 6049 GTTCTCTCGCTTATTAATTAATTTGGCAAGAAATCTGGGGGGGATTAATGAAA 6108  
 QY 301 CTTTGGGCGGCGCAACAGTGAATGTTCTGTTGCCATCGACTTGAACACCATTTGCATG 360  
 DB 6109 CTTTGGGCGGCGCAACAGTGAATGTTCTGTTGCCATCGACTTGAACACCATTTGCATG 6168  
 QY 361 GATTTCTCGAAGCACTAATCAAGCCCTGTTTGTGAGCGGTAAAGATGTGCA 420  
 DB 6169 GATTTCTCGAAGCACTAATCAAGCCCTGTTTGTGAGCGGTAAAGATGTGCA 6228  
 QY 421 CGGATGATCTTAATGATTTGCGGCACTTCAATCTTAAGACTGCGGCTTTGCAAC 480  
 DB 6229 CGGATGATCTTAATGATTTGCGGCACTTCAATCTTAAGACTGCGGCTTTGCAAC 6288  
 QY 481 CTGCAATGCGCGTGTAGAGAAATGCGAGTGAATGCGGTTTACCATTAAC 540  
 DB 6289 CTGCAATGCGCGTGTAGAGAAATGCGAGTGAATGCGGTTTACCATTAAC 6348  
 QY 541 CTTGAAGGGGCGGTGCAATTTGACATGTTAGCTTAATGTTGAAAACCTGATGCTGAAGGTT 600  
 DB 6349 CTTGAAGGGGCGGTGCAATTTGACATGTTAGCTTAATGTTGAAAACCTGATGCTGAAGGTT 6408  
 QY 601 TCGGCAAGGTTCTTCCCAACAATCGACTTGTAGACTGAGACGCTTTTGAACCA 660  
 DB 6409 TCGGCAAGGTTCTTCCCAACAATCGACTTGTAGACTGAGACGCTTTTGAACCA 6468  
 QY 661 TGTTCGATTAAGTGAAGATCGGCTTCTTCCGAGAGATTTCTTAAGCGGAAGTGGCG 720  
 DB 6469 TGTTCGATTAAGTGAAGATCGGCTTCTTCCGAGAGATTTCTTAAGCGGAAGTGGCG 6528  
 QY 721 GTCAATGCGCGTGGCAATTTCCGAGCTGTGTGGAACGAACCTGCTTCACTGCTGGGTTA 780  
 DB 6529 GTCAATGCGCGTGGCAATTTCCGAGCTGTGTGGAACGAACCTGCTTCACTGCTGGGTTA 6588

QY 781 GACGATGTTACATATATGAAGCAAGTGAATGCTGTGGAGGCAAGCTTTGTCACATGCT 840  
 Db 6589 GACATGTTACATATATGAAGCAAGTGAATGCTGTGGAGGCAAGCTTTGTCACATGCT 6648  
 QY 841 TTACGAGGACGCTCTAGTGTGTGGCGAATGGGGGCGATGCGATTTCTCTGCTGCA 900  
 Db 6649 TTACGAGGACGCTCTAGTGTGTGGCGAATGGGGGCGATGCGATTTCTCTGCTGCA 6708  
 QY 901 TTCTGCTGTTTCTTCTCTGACCGGTTACCGGCTGTCTTGATGAGGCGGTTCCCAAT 960  
 Db 6709 TTCTGCTGTTTCTTCTCTGACCGGTTACCGGCTGTCTTGATGAGGCGGTTCCCAAT 6768  
 QY 961 CCCGCGACAGTGCACACTTACTTGTCTACCAAGCGCTGCATAATCATGTGGAAGCCGG 1020  
 Db 6769 CCCGCGACAGTGCACACTTACTTGTCTACCAAGCGCTGCATAATCATGTGGAAGCCGG 6828  
 QY 1021 CAGGTGCACCGAAGCTTCCATCGGCTTACCAAGGTTGCGGTGCTTTGAAGAC 1080  
 Db 6829 CAGGTGCACCGAAGCTTCCATCGGCTTACCAAGGTTGCGGTGCTTTGAAGAC 6888  
 QY 1081 GGTTCATGAGCAGATATTTGTGTGGCTTCCGCTGCTATTAATCAAGGCTTGAAA 1140  
 Db 6889 GGTTCATGAGCAGATATTTGTGTGGCTTCCGCTGCTATTAATCAAGGCTTGAAA 6948  
 QY 1141 TCAGGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGCTGAACCGTTTCGGAGG 1200  
 Db 6949 TCAGGACACATTAAGTGGGCTCATGACTCTGCGCAATTTGCTGAACCGTTTCGGAGG 7008  
 QY 1201 GAGTCTCTTCTTCAAGGATAGAGAGATCTTTCTGGGCAACATCTCTCTGTGT- AA 1259  
 Db 7009 GAGTCTCTTCTTCAAGGATAGAGAGATCTTTCTGGGCAACATCTCTCTGTGTGAA 7068  
 QY 1260 ACATGAGTTTCTCTCATGATTTGGGACCTATTCAAGCTTAATGGGAATAGATCTGGCGG 1319  
 Db 7069 ACATGAGTTTCTCTCATGATTTGGGACCTATTCAAGCTTAATGGGAATAGATCTGGCGG 7127  
 QY 1320 GTTTGCTCAGTTTGGAAAAGCGGTTTATTAAGATCTCCGCTGGTCAACGGA 1379  
 Db 7128 GTTTGCTCAGTTTGGAAAAGCGGTTTATTAAGATCTCCGCTGGTCAACGGA 7187  
 QY 1380 TGAAGAAATCAGCGGATGTGCTTGAAGAAATCTCAAACTTCCAGCTGGAATCGCATC 1439  
 Db 7188 TGAAGAAATCAGCGGATGTGCTTGAAGAAATCTCAAACTTCCAGCTGGAATCGCATC 7247  
 QY 1440 TGAAGTGTAAAGGTGTCTGTGAGCCAGCGCATATGTCATGTTCAATCGAGCGGAT 1499  
 Db 7248 TGAAGTGTAAAGGTGTGTGTGTGAGCCAGCGCATATGTCATGTTCAATCGAGCGGAT 7307  
 QY 1500 TCAGAAAGAAAGACAAAATAAAGATTAAGGCTTAAAGAGCGGATATCTGAATTTATGA 1559  
 Db 7308 TCAGAAAGAAAGACAAAATAAAGATTAAGGCTTAAAGAGCGGATATCTGAATTTATGA 7367  
 QY 1560 TAAAGTGTGTACATCTGAGCTCGCAATATCCAACTCAGGCAATTCCTGACATGCA 1619  
 Db 7368 TAAAGTGTGTACATCTGAGCTCGCAATATCCAACTCAGGCAATTCCTGACATGCA 7427  
 QY 1620 TACCAATTTTTTCAAGCAACAGTGAACCAAGGTTGAATAAGCAACATATGACAGATC 1679  
 Db 7428 TACCAATTTTTTCAAGCAACAGTGAACCAAGGTTGAATAAGCAACATATGACAGATC 7487  
 QY 1680 GTCAAAACTTCTCTGATGACTGAAGAAATTTCTGTTAGACATATCTCCGCTTG 1739  
 Db 7488 GTCAAAACTTCTCTGATGACTGAAGAAATTTCTGTTAGACATATCTCCGCTTG 7547  
 QY 1740 TGTCTCATGAGCGGATGCAAAAGCATGTAATGCTCTGAGCTATGAGTCCGAGATCC 1799  
 Db 7548 TGTCTCATGAGCGGATGCAAAAGCATGTAATGCTCTGAGCTATGAGTCCGAGATCC 7607  
 QY 1800 GAATGTTAAAGCTTATGCTCATCACTTAATATGAGGAGAGATCCCAAGCTGTT 1859  
 Db 7608 GAATGTTAAAGCTTATGCTCATCACTTAATATGAGGAGAGATCCCAAGCTGTT 7667  
 QY 1860 GGGCGTCCCCGACAAAAGAGCATTAATGCTGTGCGGAGACGCAATTCAGATCTTT 1919

Db 7668 GGGCGTCCCCGACAAAAGAGCATTAATGCTGTGCGGAGCGCAATTCGAGATCTTT 7727  
 QY 1920 CCCGCGCTTGGCCAGACCTATTTCTGCTGCGCTGATTAAGACCAAAATGTTATCA 1979  
 Db 7728 CCCGCGCTTGGCCAGACCTATTTCTGCTGCGCTGATTAAGACCAAAATGTTATCA 7787  
 QY 1980 ACATGATGCTTACAGACGAAATGCGGGGAGCTTTCAACTCAACGGGCTGCTGA 2039  
 Db 7788 ACATGATGCTTACAGACGAAATGCGGGGAGCTTTCAAACTCAACGGGCTGCTGA 7847  
 QY 2040 GAAATTTTATTTGAAAGACTTTTCTTCAAGCACTGACACCGCTAATGATACCGAGT 2099  
 Db 7848 GAAATTTTATTTGAAAGACTTTTCTTCAAGCACTGACACCGCTAATGATACCGAGT 7907  
 QY 2100 TTACTTGGCGGGTTCAGATTTCTTCTTCAAGGTTGATGGGTGAGGCTCTATTACAG 2159  
 Db 7908 TTACTTGGCGGGTTCAGATTTCTTCTTCAAGGTTGATGGGTGAGGCTCTATTACAG 7967  
 QY 2160 CGGCTGAACGGCGTCTGTGCAATTAATCCATTTGTGAGGCAATTTGGCAAGGCAA 2219  
 Db 7968 CGGCTGAACGGCGTCTGTGCAATTAATCCATTTGTGAGGCAATTTGGCAAGGCAA 8027  
 QY 2220 TCCTCTGAAACACTCTTGGAAAGATATACTACCGCAGTGAATTAAG 2268  
 Db 8028 TCCTCTGAAACACTCTTGGAAAGATATACTACCGCAGTGAATTAAG 8076

RESULT 3  
 E00404  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Agrobacterium tumefaciens (Rhizobium radiobacter)  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 1 (bases 1 to 24595)  
 Garil, E.D., Denisu, D.S. and Richiyado, E.B.  
 SELECTION USING OPINE SYNTHASE GENE  
 Patent: JP 1985156333-A 1 16-AUG-1985;  
 AGURIJENETIKUSU RES ASSOC LTD  
 OS Agrobacterium tumefaciens  
 PN JP 1985156333-A/1  
 PD 16-AUG-1985  
 PF 14-SEP-1984 JP 1984193841  
 PR 14-SEP-1983 US 83 532280  
 PI GARIL, E.B. DABRU, DENISU DABURITU SATSUTON,  
 RI RICHIIYADO EFU BEIRKA  
 PC A01H1/00, C12N5/00, C12N15/00, C12N5/00, C12R1/91; CC  
 strandedness: Double;  
 CC topology: Linear;  
 CC \*source: strain=15955;  
 FH key  
 FH location/Qualifiers  
 FT repeat\_unit  
 FT /note='ROTL(A)',  
 FT repeat\_unit  
 FT /note='ROTL(B)',  
 FT repeat\_unit  
 FT /note='ROTL(C)',  
 FT repeat\_unit  
 FT /note='ROTL(D)',  
 FT CDS  
 FT /product='octopine synthase',  
 FT CDS  
 FT /product='agropine synthase',  
 FT CDS  
 FT /product='agropine synthase',  
 FT CDS  
 FT /product='mannopine synthase'.

FEATURES  
source 1. .24595  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:358"  
BASE COUNT 6534 a 5509 c 5793 g 6759 t  
ORIGIN  
Query Match 98.4%; Score 2230.6; DB 6; Length 24595;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;  
QY 1 ATGTCACTTCACTCTCTTGATTAACAAGTGCATCATCTCCCAACCAAAATGTGGAT 60  
Db 5809 ATGTCACTTCACTCTCTTGATTAACAAGTGCATCATCTCCCAACCAAAATGTGGAT 5868  
QY 61 CTGACAAATGGTGAATGAAGCGGATGAATGGACCGCAGGGTTTCCGATGCTTTAGAA 120  
Db 5869 CTGACAAATGGTGAATGAAGCGGATGAATGGACCGCAGGGTTTCCGATGCTTTAGAA 5928  
QY 121 CGAAGAGCTTCTAGGGGAGAGAGATTACCAAGCTCCACCGAGTGCAGCGCTGGGTTA 180  
Db 5929 CGAAGAGCTTCTAGGGGAGAGAGATTACCAAGCTCCACCGAGTGCAGCGCTGGGTTA 5988  
QY 181 GCTTGAAGAGCTGGCCGATGGTCGCTTCCCGAGATCTCAGCTGGTGAAGATGCA 240  
Db 5989 GCTTGAAGAGCTGGCCGATGGTCGCTTCCCGAGATCTCAGCTGGTGAAGATGCA 6048  
QY 241 GTTCTCTCCGCTTATATCTATATTTGGCAAGAAATCTGGGCGGATCTTGAATGGAA 300  
Db 6049 GTTCTCTCCGCTTATATCTATATTTGGCAAGAAATCTGGGCGGATCTTGAATGGAA 6108  
QY 301 CCTTGGGCGGCGGACAGTGAAGTGTCTGTTGCCATCGACTTGGACACATTTGGATG 360  
Db 6109 CCTTGGGCGGCGGACAGTGAAGTGTCTGTTGCCATCGACTTGGACACATTTGGATG 6168  
QY 361 GATTTCCTCGAAGACAACTAATCCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 420  
Db 6169 GATTTCCTCGAAGACAACTAATCCAAGCCCTGTTTTGCTGAGCGGTAAAGATGCA 6228  
QY 421 CCGATTGATCTTAGTCAATTTGTTGGCCATTTCATCTTAAGACTGCGGCTTTTCCAA 480  
Db 6229 CCGATTGATCTTAGTCAATTTGTTGGCCATTTCATCTTAAGACTGCGGCTTTTCCAA 6288  
QY 481 CTGCGCAATGGCGCTGTACAGGAATGGACAGATGAAGACGTTACCGGTTTACCATAC 540  
Db 6289 CTGCGCAATGGCGCTGTACAGGAATGGACAGATGAAGACGTTACCGGTTTACCATAC 6348  
QY 541 CTGGAAGGGGCGGTGCCATTGACATGTAGCTTATGTCGAAACCTGATGCTGAAGGT 600  
Db 6349 CTGGAAGGGGCGGTGCCATTGACATGTAGCTTATGTCGAAACCTGATGCTGAAGGT 6408  
QY 601 TCGGAGAGTTCTTTCCAAATCGACTTCTGACGATGACAGCCGTTTTTGGACAA 660  
Db 6409 TCGGAGAGTTCTTTCCAAATCGACTTCTGACGATGACAGCCGTTTTTGGACAA 6468  
QY 661 TGTTCGATAGTGAACGATCGGCTTTTCCGAGAGATTTCTTAAGCCGAAAGTGGCG 720  
Db 6469 TGTTCGATAGTGAACGATCGGCTTTTCCGAGAGATTTCTTAAGCCGAAAGTGGCG 6528  
QY 721 GTGATGGCGCTGGCATTTCCGCACTGCTGATGGCAACGATCTTCACTGCTGGGTA 780  
Db 6529 GTGATGGCGCTGGCATTTCCGCACTGCTGATGGCAACGATCTTCACTGCTGGGTA 6588  
QY 781 GACGATGTTACATATATGAAGCAAGTATGTTGTTGAGGCAAGCTTTGATCATGCT 840  
Db 6589 GACGATGTTACATATATGAAGCAAGTATGTTGTTGAGGCAAGCTTTGATCATGCT 6648  
QY 841 TTCAAGGAGCGCTCTATGTCTGTGGCCGAATGGGGCGATGCAATTTCTCTGCTGCA 900  
Db 6649 TTCAAGGAGCGCTCTATGTCTGTGGCCGAATGGGGCGATGCAATTTCTCTGCTGCA 6708  
QY 901 TTCTGCTGTTTCTTCTCTGACGCTTACGCGCTGTCTTGATGAGCGCTTCCCAAT 960

Db 6709 TTCTGCTGTTTCTTCTCTGACGCTTACGCGCTGTCTTGATGAGCGCTTCCCAAT 6768  
QY 961 CCCGACAGTGCACACTTACTGTGTACCAAGGCGTCCAAATACATGTGAAAGCGGG 1020  
Db 6769 CCCGACAGTGCACACTTACTGTGTGTACCAAGGCGTCCAAATACATGTGAAAGCGGG 6828  
QY 1021 CAGCTGCCACCGAAGCTGTTCCATCGCTTTTCAACCGTTGCGCTGCTTTGAAAGAC 1080  
Db 6829 CAGCTGCCACCGAAGCTGTTCCATCGCTTTTCAACCGTTGCGCTGCTTTGAAAGAC 6888  
QY 1081 GGTTCATGACGAGATATTTGTTGGCTTGGCTGTGCTTATTAAGGCTTTGAA 1140  
Db 6889 GGTTCATGACGAGATATTTGTTGGCTTGGCTGTGCTTATTAAGGCTTTGAA 6948  
QY 1141 TCAGGACATATGAGGCGCTCATGCTCTGGCAAAATTTGGGTGAACGTTTGGGAG 1200  
Db 6949 TCAGGACATATGAGGCGCTCATGCTCTGGCAAAATTTGGGTGAACGTTTGGGAG 7008  
QY 1201 GAGTCTCTCTTCAAGGATAGAGAGATCTTCTGGGACACATCTCTGCTGTGT -AA 1259  
Db 7009 GAGTCTCTCTTCAAGGATAGAGAGATCTTCTGGGACACATCTCTGCTGTGTGAA 7068  
QY 1260 ACATGAGTTTCTCTCATGATTTGGACCTTATCAAGCTAATGGGAATAGATCTGGCG 1319  
Db 7069 ACATGAGTTTCTCTCATGATTTGGACCTTATCAAGCTAATGGGAATAGATCTGGCG 7127  
QY 1320 GTTGGTCCAGTTTGTGAAGGGGTTTATGAGATCTCCGCTTGGTATCAAGGATA 1379  
Db 7128 GTTGGTCCAGTTTGTGAAGGGGTTTATGAGATCTCCGCTTGGTATCAAGGATA 7187  
QY 1380 TGAAGAAATACAGCGATGTGCTTGAGAAATCTCAGAACTTCCAGTGCATCTC 1439  
Db 7188 TGAAGAAATACAGCGATGTGCTTGAGAAATCTCAGAACTTCCAGTGCATCTC 7247  
QY 1440 TGAAGTGTAAAGGTGTGTCTGTGACCAAGCGCATATGCCATGTTCAAGTCAAGGCGAT 1499  
Db 7248 TGAAGTGTAAAGGTGTGTCTGTGACCAAGCGCATATGCCATGTTCAAGTCAAGGCGAT 7307  
QY 1500 TCGAAGGAAAGACAAATAAAGTAAAGGCTTAAGAGGGGATATCTGAATTTATGA 1559  
Db 7308 TCGAAGGAAAGACAAATAAAGTAAAGGCTTAAGAGGGGATATCTGAATTTATGA 7367  
QY 1560 TAAAGTGTGTGTCACATCTGGAATCGCAATATCAACTCAGGACTTGCCTGACATGCA 1619  
Db 7368 TAAAGTGTGTGTCACATCTGGAATCGCAATATCAACTCAGGACTTGCCTGACATGCA 7427  
QY 1620 TACCAATATTTTACGACACAGTGAACCAAGCGTTGATTAAGCCATATGACAGATC 1679  
Db 7428 TACCAATATTTTACGACACAGTGAACCAAGCGTTGATTAAGCCATATGACAGATC 7487  
QY 1680 GTCAAAATCTTCTGATGATGGAAGAAATTTCTGGTTAGACCATATCTCCGCTTG 1739  
Db 7488 GTCAAAATCTTCTGATGATGGAAGAAATTTCTGGTTAGACCATATCTCCGCTTG 7547  
QY 1740 TGTCTCATGAGAGGATCGCAAAACAGTGTATCTCTGACATATGAGTGCAGAGATCC 1799  
Db 7548 TGTCTCATGAGAGGATCGCAAAACAGTGTATCTCTGACATATGAGTGCAGAGATCC 7607  
QY 1800 GAAATGTAAGTCTAGTGTCTCATGATTAATGAGAGAGATCTCCCAAGCTGT 1859  
Db 7608 GAAATGTAAGTCTAGTGTCTCATGATTAATGAGAGAGATCTCCCAAGCTGT 7667  
QY 1860 GGGGTCCTCCGCAAAAAGAGGATTAATGTCGCTGGGAGGCAATTTGAGATCTT 1919  
Db 7668 GGGGTCCTCCGCAAAAAGAGGATTAATGTCGCTGGGAGGCAATTTGAGATCTT 7727  
QY 1920 CCCGCGTTTCCACACACTATTTCTGCTGCTGCTGATTAAGCAACCAATGTTATCA 1979  
Db 7728 CCCGCGTTTCCACACACTATTTCTGCTGCTGCTGATTAAGCAACCAATGTTATCA 7787  
QY 1980 ACATGATGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACGCGGTGTGA 2039

Db 7788 ACATGATGGCTTAACAGACGAAATGCGGGAGCTTTCAACTCAACCGCGCTGTGA 7847  
Qy 2040 GGAATTTTATTTGAAAGACTTTTCTTTCAAGCACTGAGACGGCTAATGATACCGGAGT 2099  
Db 7848 GGAATTTTATTTGAAAGACTTTTCTTTCAAGCACTGAGACGGCTAATGATACCGGAGT 7907  
Qy 2100 TTACTTGGCGGGTTGCAAGTTGTTCTTCAAGGTGATGGGTGAGGGTCTATTTCAGAC 2159  
Db 7908 TTACTTGGCGGGTTGCAAGTTGTTCTTCAAGGTGATGGGTGAGGGTCTATTTCAGAC 7967  
Qy 2160 CGCGTGTAAAGCGCGTCTGTGCAATTATCCACAATTGTGAGGCAATTTGGCAAGGCGCA 2219  
Db 7968 CGCGTGTAAAGCGCGTCTGTGCAATTATCCACAATTGTGAGGCAATTTGGCAAGGCGCA 8027  
Qy 2220 TCCTCTGGAACACTCTTGAAGATATTACTACCGAGTAATAATTAG 2268  
Db 8028 TCCTCTGGAACACTCTTGAAGATATTACTACCGAGTAATAATTAG 8076

RESULT 4  
LOCUS E00546 24595 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA fragment of T-DNA of Ti-plasmid, pTI15955.  
ACCESSION E00546  
VERSION E00546.1 GI:2168825  
KEYWORDS JP 1985256383-A/1.  
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)  
ORGANISM Agrobacterium tumefaciens  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
1 (bases 1 to 24595)  
Richiyoado, E.B. and Jiyon, D.K.  
PROMOTOR OF OCTOPINE T-DNA  
Patent: JP 1985256383-A 1 18-DEC-1985;  
AGURIJENETIKUSU RES ASSOC LTD  
OS Agrobacterium tumefaciens  
PN JP 1985256383-A/1  
PD 18-DEC-1985  
PE 18-NOV-1984 JP 1984244307  
PR 18-NOV-1983 US 83 553786  
PI RICHIVADO EBU BEIKAA, JIYON DEI KENPU  
PC C12N15/00, A01H1/00, C12N1/20, C12N5/00//C07H21/04, (C12N1/20, PC  
C12R1/01),  
PC (C12N5/00, C12R1/91);  
CC strandedness: Single;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key Location/Qualifiers  
FH  
FT CDS >1..<24595  
FT /gene='tme'  
FT >1..<24595  
FT /gene='tme'  
FT >1..<24595  
FT /gene='tme'  
FT >1..<24595  
FT /gene='tme'  
FT CDS >1..<24595  
FT /gene='tme'  
FT >1..<24595  
FT /gene='tme'  
FT CDS >1..<24595  
FT /gene='tme'  
FT >1..<24595  
FT /gene='tme'

FEATURES  
source 1..24595  
location/Qualifiers  
1..24595  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:358"

BASE COUNT 6534 a 5509 c 5793 g 6759 t

ORIGIN  
Query Match 98.4%; Score 2230.6; DB 6; Length 24595;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;  
Qy 1 ATGTGAGCTTACCTCTCTTGAATACAGTGCATCTCCCAACCAAAATGTGAT 60  
|||||

Db 5809 ATGTGAGCTTACCTCTCTTGAATACAGTGCATCTCCCAACCAAAATGTGAT 5868  
Qy 61 CTGCAATGCTGATTAAGGCGATGTAATTGACACCGAGGTTTCCATGCTTTAGAA 120  
Db 5869 CTGCAATGCTGATTAAGGCGATGTAATTGACACCGAGGTTTCCATGCTTTAGAA 5928  
Qy 121 CGAAGAGCTTCTAGGGAAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGTTA 180  
Db 5929 CGAAGAGCTTCTAGGGAAGAGAGATTACTCAAGCTCCACGAGTGCAGCGCTGGTTA 5988  
Qy 181 GCTTGCAAAAGGCTGCGCATGCTGCTTCCGAGATCTCAGCTGTGAAAGTAGCA 240  
Db 5989 GCTTGCAAAAGGCTGCGCATGCTGCTTCCGAGATCTCAGCTGTGAAAGTAGCA 6048  
Qy 241 GTTCTCTCCGCTTATATCTATATGCAAGAAATTTCTGGGCGGATCTTGAATGAAA 300  
Db 6049 GTTCTCTCCGCTTATATCTATATGCAAGAAATTTCTGGGCGGATCTTGAATGAAA 6108  
Qy 301 CTTTGGGCGGCGCAACAGTGAAGTGTCTGTTGCCATCGACTTTGACACATTTGCATG 360  
Db 6109 CTTTGGGCGGCGCAACAGTGAAGTGTCTGTTGCCATCGACTTTGACACATTTGCATG 6168  
Qy 361 GATTTCTCCGAAGCACAACTAATCCAAGCCTGTTTGTGACGCGTAAAGATGTGA 420  
Db 6169 GATTTCTCCGAAGCACAACTAATCCAAGCCTGTTTGTGACGCGTAAAGATGTGA 6228  
Qy 421 CGATTTGATCTTATGATTTTGTGCGCAATTTCAATCTTAAGCTGCGGCTTTGCAAC 480  
Db 6229 CGATTTGATCTTATGATTTTGTGCGCAATTTCAATCTTAAGCTGCGGCTTTGCAAC 6288  
Qy 481 CTGCAATGCGCGCTGACGAAATGCAAGTGAATGCGGCTTTACATTAAC 540  
Db 6289 CTGCAATGCGCGCTGACGAAATGCAAGTGAATGCGGCTTTACATTAAC 6348  
Qy 541 CTTGAAAGGCGCGCTGCAATTTGACATGTAATGTTGCAAACTGATGCTGAAGGCT 600  
Db 6349 CTTGAAAGGCGCGCTGCAATTTGACATGTAATGTTGCAAACTGATGCTGAAGGCT 6408  
Qy 601 TCGGCAAGGCTTCTTTCAACATGCACTGCTTACGACCTGCAACCGTTTGTACCA 660  
Db 6409 TCGGCAAGGCTTCTTTCAACATGCACTGCTTACGACCTGCAACCGTTTGTACCA 6468  
Qy 661 TGTTCGATAGTGAAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGGCGG 720  
Db 6469 TGTTCGATAGTGAAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGGCGG 6528  
Qy 721 GTCAATGGCGTGGCAATTTCCGACTCGTGTGGCAAGCAACTGCTTATGCTGGGTA 780  
Db 6529 GTCAATGGCGTGGCAATTTCCGACTCGTGTGGCAAGCAACTGCTTATGCTGGGTA 6588  
Qy 781 GACGATGTTACATATATGAAGCAAGTGAATCGTGTGAGCGAAGCTTTGTCATGCT 840  
Db 6589 GACGATGTTACATATATGAAGCAAGTGAATCGTGTGAGCGAAGCTTTGTCATGCT 6648  
Qy 841 TTCAGGAGGCTTCTTAATGTCGAGCGCAAAATGGGCGCATGCAATTTCTCTGCTGCA 900  
Db 6649 TTCAGGAGGCTTCTTAATGTCGAGCGCAAAATGGGCGCATGCAATTTCTCTGCTGCA 6708  
Qy 901 TTCTGCTTTTCTTCTCTGAGCGTTACGAGCTGTCTTCAATGAGCGCTTCCCAAT 960  
Db 6709 TTCTGCTTTTCTTCTCTGAGCGTTACGAGCTGTCTTCAATGAGCGCTTCCCAAT 6768  
Qy 961 CCGGCGCAAGTGAAGCTTCTGATGTCACCAAGCGTCCAAATACATGTGAAAGCGG 1020  
Db 6769 CCGGCGCAAGTGAAGCTTCTGATGTCACCAAGCGTCCAAATACATGTGAAAGCGG 6828  
Qy 1021 CAGCTGCAAGTGAAGCTTCTGATGTCACCAAGCGTCCAAATACATGTGAAAGCG 1080  
Db 6829 CAGCTGCAAGTGAAGCTTCTGATGTCACCAAGCGTCCAAATACATGTGAAAGCG 6888  
Qy 1081 GGTTCATGAGCGAGATATGTTGTTGCTGCTGCTGCTGCTTACTAGGCTTTGAA 1140  
Db 6889 GGTTCATGAGCGAGATATGTTGTTGCTGCTGCTGCTGCTTACTAGGCTTTGAA 6948

QY 1141 TCAGACACATTAGTGGGCTCATGACTCTGGCAATTGCTGAACGGTTTGGGAGG 1200  
 DB 6949 TCAGAGACATTAGTGGGCTCATGACTCTGGCAATTGCTGAACGGTTTGGGAGG 7008  
 QY 1201 GAGTCTTCTCTCAGGAGATAGAGAGATCTTCTGGGCAACATCTCTCTGGTGT-AA 1259  
 DB 7009 GAGTCTTCTCTCAGGAGATAGAGAGATCTTCTGGGCAACATCTCTCTGGTGTGAA 7068  
 QY 1260 ACATGAGTTTCTCATGATTGGAGCTTTCAGCTAATGGGAATAGATCTGGCGGG 1319  
 DB 7069 ACATGAGTTTCTCATGATTGGAGCTTTCAGCTAATGGGAATAGATCTGGCG-GG 7127  
 QY 1320 GTTTGGCCAGTTTGGAAAGCGGGTTTATGAGATCTCGGCTGGTGCATCAACGGATA 1379  
 DB 7128 GTTTGGCCAGTTTGGAAAGCGGGTTTATGAGATCTCGGCTGGTGCATCAACGGATA 7187  
 QY 1380 TGAAGAAATCAGCGGATGTGCTGAGGATCTCAGAACTTCCAGCTCGATCCGATC 1439  
 DB 7188 TGAAGAAATCAGCGGATGTGCTGAGGATCTCAGAACTTCCAGCTCGATCCGATC 7247  
 QY 1440 TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAAGCGCAT 1499  
 DB 7248 TGAAGTGTAAACGGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAAGCGCAT 7307  
 QY 1500 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAGAGGGGATATCTGAACCTTTATGA 1559  
 DB 7308 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAGAGGGGATATCTGAACCTTTATGA 7367  
 QY 1560 TAAAGTGTGTGATCAGTCTGAGCTCGCAATATCCAACTCAGGCAATTCCTGATCCGA 1619  
 DB 7368 TAAAGTGTGTGATCAGTCTGAGCTCGCAATATCCAACTCAGGCAATTCCTGATCCGA 7427  
 QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 1679  
 DB 7428 TACCAATATTTTTCAGGACCAAGTGAACCAAGCGGTTGATTAACGCCATATGACAGATC 7487  
 QY 1680 GTCAAAACCTTCTCTGATGATGAGGAAATCTGGTTAGACCATATCTCTCCGCTTGG 1739  
 DB 7488 GTCAAAACCTTCTCTGATGATGAGGAAATCTGGTTAGACCATATCTCTCCGCTTGG 7547  
 QY 1740 TGTCTCATGAGAGGATTCGCAAAAGCAGTGTATTCCTGAGTATGATGCGAGATCC 1799  
 DB 7548 TGTCTCATGAGAGGATTCGCAAAAGCAGTGTATTCCTGAGTATGATGCGAGATCC 7667  
 QY 1800 GAATGTAAAGTCTAGTCTCATCACTTAATACATGGAGAGACATCTCCCAAGCTGTT 1859  
 DB 7608 GAATGTAAAGTCTAGTCTCATCACTTAATACATGGAGAGACATCTCCCAAGCTGTT 7667  
 QY 1860 GGGGGTCCCGCAAAAAGAGGATATGCTGCTGCGGAGGAGCATTTTGAATCTTT 1919  
 DB 7668 GGGGGTCCCGCAAAAAGAGGATATGCTGCTGCGGAGGAGCATTTTGAATCTTT 7727  
 QY 1920 CCGGCGCTTGGCCAGCACTATTTCTGCTGCTGCTGATTAACGCAAAATGTTATCA 1979  
 DB 7728 CCGGCGCTTGGCCAGCACTATTTCTGCTGCTGCTGATTAACGCAAAATGTTATCA 7787  
 QY 1980 ACATGATGCTTACAGACAGAAATGCGGGGAGCTTTCAAACTCAACGGCGCTGTGA 2039  
 DB 7788 ACATGATGCTTACAGACAGAAATGCGGGGAGCTTTCAAACTCAACGGCGCTGTGA 7847  
 QY 2040 GGAATTTTATTCGAAGAATCTTTCTTCAAGCACTGAGACACGGCTAATGATACCGGAGT 2099  
 DB 7848 GGAATTTTATTCGAAGAATCTTTCTTCAAGCACTGAGACACGGCTAATGATACCGGAGT 7907  
 QY 2100 TTAATTTGGCGGTTGACAGTTTCTTCAAGTGTGATGGTGGAGGCTTATTCAGAC 2159  
 DB 7908 TTAATTTGGCGGTTGACAGTTTCTTCAAGTGTGATGGTGGAGGCTTATTCAGAC 7967  
 QY 2160 CGGCTGTAAACGCGCTGTGTCAATTATCACAATTTGAGAGGCAATTTGGCAAGGCGCA 2219  
 DB 7968 CGGCTGTAAACGCGCTGTGTCAATTATCACAATTTGAGAGGCAATTTGGCAAGGCGCA 8027

QY 2220 TCCTCTCGAACAACCTTTGAGAGATATTAACCGCAGTACGAAATTTAG 2268  
 DB 8028 TCCTCTCGAACAACCTTTGAGAGATATTAACCGCAGTACGAAATTTAG 8076  
 RESULT 5  
 AF242881  
 LOCUS  
 DEFINITION  
 Agrobacterium tumefaciens octopine-type Ti plasmid, complete  
 ACCESSION  
 AF242881 194140 bp DNA circular BCT 26-JUN-2000  
 VERSION  
 AF242881.1 GI:8572673  
 KEYWORDS  
 Agrobacterium tumefaciens (Rhizobium radiobacter)  
 SOURCE  
 ORGANISM  
 Agrobacterium tumefaciens  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 1 (bases 193234 to 194140; 1 to 23692)  
 Barker, R.F., Idler, K.B., Thompson, D.V. and Kemp, J.D.  
 Nucleotide sequence of the T-DNA region from the Agrobacterium  
 tumefaciens octopine Ti plasmid pTi1595  
 Plant Mol. Biol. 2, 335-350 (1983)  
 2 (bases 16643 to 168193)  
 original sequence presented in GenBank Accession Number X00493  
 Yanofsky, M.F. and Nester, E.W.  
 Molecular characterization of a host-range-determining locus from  
 Agrobacterium tumefaciens  
 J. Bacteriol. 168 (1), 244-250 (1986)  
 3 (bases 168089 to 173951)  
 Yanofsky, M.F., Porter, S.G., Young, C., Albright, L.M., Gordon, M.P.  
 and Nester, E.W.  
 The virD operon of Agrobacterium tumefaciens encodes a  
 site-specific endonuclease  
 Cell 47 (3), 471-477 (1986)  
 4 (bases 175278 to 177719)  
 Winans, S.C., Allenza, P., Stachel, S.E., McBride, K.E. and Nester, E.W.  
 Characterization of the virE operon of the Agrobacterium Ti plasmid  
 pTiA6  
 Nucleic Acids Res. 15 (2), 825-837 (1987)  
 5 (bases 50125 to 54471)  
 Habeeb, L.F., Wang, L. and Winans, S.C.  
 Transcription of the octopine catabolism operon of the  
 Agrobacterium tumor-inducing plasmid pTiA6 is activated by a  
 lyxR-type regulatory protein  
 Mol. Plant Microbe Interact. 4 (4), 379-385 (1991)  
 6 (bases 50125 to 54471)  
 Valdivia, R.H., Wang, L. and Winans, S.C.  
 Characterization of a putative periplasmic transport system for  
 octopine accumulation encoded by Agrobacterium tumefaciens Ti  
 plasmid pTiA6  
 J. Bacteriol. 173 (20), 6398-6405 (1991)  
 7 (bases 68024 to 76460)  
 Kim, K.S. and Fairhead, S.K.  
 Ti plasmid-encoded genes responsible for catabolism of the crown  
 gall opine mannopine by Agrobacterium tumefaciens are homologs of  
 the T-region genes responsible for synthesis of this opine by the  
 plant tumor  
 J. Bacteriol. 178 (11), 3275-3284 (1996)  
 MEDLINE  
 PUBMED  
 8 (bases 39167 to 45466)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE

- AUTHORS  
TITLE  
Fugua, C. and Winans, S.C.  
Localization of OccR-activated and Trar-activated promoters that  
express two ABC-type permeases and the trar gene of Ti plasmid  
pTiR10
- JOURNAL  
MEDLINE  
PUBMED  
96405643  
88097772  
9 (bases 22855 to 39243; 96353 to 111409)  
Alt-Morbe, J., Stryker, J.L., Fugua, C., Li, P.L., Farrand, S.K. and  
Winans, S.C.  
The conjugal transfer system of Agrobacterium tumefaciens  
octopine-type Ti plasmids is closely related to the transfer system  
of an IncP plasmid and distantly related to Ti plasmid vir genes  
J. Bacteriol. 178 (14), 4248-4257 (1996)
- JOURNAL  
MEDLINE  
PUBMED  
8763954  
10 (bases 54312 to 62806)  
Oger, P., Kim, K.-S., Sackett, R.L., Piper, K.R. and Farrand, S.K.  
Octopine-type Ti plasmids code for a mannopine-inducible  
dominant-negative allele of trar, the quorum-sensing activator that  
regulates Ti plasmid conjugal transfer  
Mol. Microbiol. 27 (2), 277-288 (1998)
- JOURNAL  
MEDLINE  
PUBMED  
9484884  
11 (bases 133963 to 138972)  
Kalogeraki, V.S. and Winans, S.C.  
Wound-released chemical signals may elicit multiple responses from  
an Agrobacterium tumefaciens strain containing an octopine-type Ti  
plasmid  
J. Bacteriol. 180 (21), 5660-5667 (1998)
- JOURNAL  
MEDLINE  
PUBMED  
99009000  
9791116  
12 (bases 76455 to 98723)  
Ly, S.M., Jafari, S. and Winans, S.C.  
Mannopine acid and agropine acid catabolism region of the  
octopine-type Ti plasmid pTi15955  
Mol. Microbiol. 31 (1), 339-347 (1999)
- JOURNAL  
MEDLINE  
PUBMED  
99141607  
9987134  
13 (bases 1 to 194140)  
Winans, S.C., Zhu, J., Oger, P.M., Schrammeyer, B., Hooykaas, P.J. and  
Farrand, S.K.  
Octopine-type Ti plasmid sequence  
Unpublished
- JOURNAL  
PUBMED  
14 (bases 68024 to 76460)  
Kim, K.-S.  
Direct Submission  
Submitted (10-JAN-1995) Kun-Soo Kim, Microbiology, University of  
Illinois at Urbana-Champaign, 1201 W. Gregory Dr., Urbana, IL  
61801, USA
- REFERENCE  
AUTHORS  
TITLE  
Winans, S.C.  
Direct Submission  
Submitted (20-DEC-1995) Stephen C. Winans, Microbiology, Cornell  
University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE  
AUTHORS  
TITLE  
Fugua, C.  
Direct Submission  
Submitted (07-FEB-1996) Clay Fugua, Biology, Trinity University,  
715 Stadium Dr., San Antonio, TX 78212, USA
- REFERENCE  
AUTHORS  
TITLE  
Kim, K.-S., Sackett, R.L. and Farrand, S.K.  
Direct Submission  
Submitted (04-JUN-1996) Crop Sciences, University of Illinois at  
Urbana-Champaign, 240 ERM, 1201 W. Gregory Dr., Urbana, IL 61801,  
USA
- REFERENCE  
AUTHORS  
TITLE  
Zhu, J. and Winans, S.C.  
Direct Submission  
Submitted (17-NOV-1997) Section of Microbiology, Cornell  
University, Wing Hall, Ithaca, NY 14853, USA
- REFERENCE  
AUTHORS  
TITLE  
Winans, S.C. and Ly, S.  
18 (bases 186239 to 193239)

TITLE  
JOURNAL  
Direct Submission  
Submitted (20-NOV-1997) Microbiology, Cornell University, Wing  
Hall, Ithaca, NY 14853, USA

REFERENCE  
AUTHORS  
TITLE  
Zhu, J. and Winans, S.C.  
Direct Submission  
Submitted (25-NOV-1997) Section of Microbiology, Cornell  
University, Wing Hall, Ithaca, NY 14853, USA

REFERENCE  
AUTHORS  
TITLE  
Stryker, J.L., Nantis, N.J., Kalogeraki, V.S. and Winans, S.C.  
Direct Submission  
Submitted (23-DEC-1997) Microbiology, Cornell University, Ithaca,  
NY 14853, USA

REFERENCE  
AUTHORS  
TITLE  
Winans, S.C.  
Direct Submission  
Submitted (03-AUG-1998) Microbiology, Cornell University, Wing  
Hall, Ithaca, NY 14853, USA

REFERENCE  
AUTHORS  
TITLE  
Oger, P.M.  
Direct Submission  
Submitted (29-DEC-1998) Crop Sciences, University of Illinois at  
Urbana-Champaign, 361 ERM, 1201, W. Gregory Dr., Urbana, IL 61801,  
USA

REFERENCE  
AUTHORS  
TITLE  
Zhu, J. and Winans, S.C.  
Direct Submission  
Submitted (18-AUG-1999) Section of Microbiology, Cornell  
University, Wing Hall, Ithaca, NY 14853, USA

REFERENCE  
AUTHORS  
TITLE  
Zhu, J., Oger, P.M., Schrammeyer, B., Hooykaas, P.J., Farrand, S.K. and  
Winans, S.C.  
Direct Submission  
Submitted (07-MAR-2000) Microbiology, Cornell University, Wing  
Hall, Ithaca, NY 14853, USA

COMMENT  
On or before Jun 20, 2000 this sequence version replaced  
gi:3377769, gi:1215729, gi:1381799, gi:797330, gi:2665704,  
gi:2749889, gi:5738260, gi:2773254, gi:2773255, gi:2773256,  
gi:154820, gi:154827, gi:2662524, gi:39122, gi:2687604, gi:154767,  
gi:154773, gi:6689862.

FEATURES  
source  
1..194140  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:358"  
/plasmid="Ti"  
/note="extrachromosomal  
octopine-type"  
13244..13618  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:358"  
/insertion\_seg="ybd4"  
13636..14181  
/organism="Agrobacterium tumefaciens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:358"  
/insertion\_seg="ybe"  
/note="similar to IS66"  
14183..14666  
/organism="Agrobacterium tumefaciens"

Query Match 98.4%; Score 2230.6; DB 1; Length 194140;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2258; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 1 ATGTACGCTTACCTCTCTGATATACCGATGATCACTCCCAACAAATGCGGAT 60  
DB 4902 ATGTACGCTTACCTCTCTTATATACCGATGATCACTCCCAACAAATGCGGAT 4961  
QY 61 CTGACAAATGCGATTAAGCGGATGAATGACCGGAGGTTCCGATCCTTTAGAA 120  
DB 4962 CTGACAAATGCGATTAAGCGGATGAATGACCGGAGGTTCCGATCCTTTAGAA 5021

QY 121 CGAAGCTTCTAGGGGAAAGAGATTACTCAAGCTCCACGAGTGCAGCGTGGTTA 180  
DB 5022 CGAAGGCTTCTAGGGGAAAGAGATTACTCAAGCTCCACGAGTGCAGCGTGGTTA 5081  
QY 181 GCTTGGCAAAAGCTGGCCGATGGTCCCTCCGAGATCTCAGCTGGTGAAGTAGCA 240  
DB 5082 GCTTGGCAAAAGCTGGCCGATGGTCCCTCCGAGATCTCAGCTGGTGAAGTAGCA 5141  
QY 241 GTTCTCTCCGCTTATATCTATATATGGCAAAAGAAATCTGGGGGGGATACTTGAATGGAA 300  
DB 5142 GTTCTCTCCGCTTATATCTATATATGGCAAAAGAAATCTGGGGGGGATACTTGAATGGAA 5201  
QY 301 CCTTGGGCGGGGCAACAGTAGTGGTCTGTGGCCATCGACTTGGACACCATTTGCAATG 360  
DB 5202 CCTTGGGCGGGGCAACAGTAGTGGTCTGTGGCCATCGACTTGGACACCATTTGCAATG 5261  
QY 361 GATTTCTCCGAGACACAACTAATCCAAAGCCCTGTCTTGTGAGCGGTAAGATGTGA 420  
DB 5262 GATTTCTCCGAGACACAACTAATCCAAAGCCCTGTCTTGTGAGCGGTAAGATGTGA 5321  
QY 421 CCGATTGATTTAGTCAATTTCTGTGGCCATTTCAATCTCTAAGACTGCCGGCTTTCCAAAC 480  
DB 5322 CCGATTGATTTAGTCAATTTCTGTGGCCATTTCAATCTCTAAGACTGCCGGCTTTCCAAAC 5381  
QY 481 CTGCGCAATGCGCGGTATGACGAAATGGCAGATGAATGCGTTACCGGGTTTACCATPACC 540  
DB 5382 CTGCGCAATGCGCGGTATGACGAAATGGCAGATGAATGCGTTACCGGGTTTACCATPACC 5441  
QY 541 CTTGAAGGGGCGGTGCCATTTGACATGTAGCTTATGTGTGAAACCTGATGCTGAAGGGT 600  
DB 5442 CTTGAAGGGGCGGTGCCATTTGACATGTAGCTTATGTGTGAAACCTGATGCTGAAGGGT 5501  
QY 601 TCGGCAAGTTCCTTTCCAAATCGAATTTGCTCTAAGACTGCGACACCGTCTTTTGAACCA 660  
DB 5502 TCGGCAAGTTCCTTTCCAAATCGAATTTGCTCTAAGACTGCGACACCGTCTTTTGAACCA 5561  
QY 661 TGTTCGATAGTGAACGAGTCGGCTTCTTCCGAGGATGTTCTTAAGCGGAAAGTGCG 720  
DB 5562 TGTTCGATAGTGAACGAGTCGGCTTCTTCCGAGGATGTTCTTAAGCGGAAAGTGCG 5621  
QY 721 GTCATGGCGCTGGCATTTCCGAGCTGTGTGTGCAACGAACTGCTTCAATGCTGGGGTA 780  
DB 5622 GTCATGGCGCTGGCATTTCCGAGCTGTGTGTGCAACGAACTGCTTCAATGCTGGGGTA 5681  
QY 781 GAGGATGTTCATATATGAAGCAAGTATGCTGTGTGAGGCAAGCTTTGTGACATGCT 840  
DB 5682 GAGGATGTTCATATATGAAGCAAGTATGCTGTGTGAGGCAAGCTTTGTGACATGCT 5741  
QY 841 TTCAAGGAGCGCTCTAGTGTGTGTGCGGAAATGGGGGCGATGCGATTTCTCTGCTGCA 900  
DB 5742 TTCAAGGAGCGCTCTAGTGTGTGTGCGGAAATGGGGGCGATGCGATTTCTCTGCTGCA 5801  
QY 901 TTCTGTCTTGTCTTCTCTGAGCGGTACGCGCTGTCTTGTGAGAGCGGCTTCCCAAT 960  
DB 5802 TTCTGTCTTGTCTTCTCTGAGCGGTACGCGCTGTCTTGTGAGAGCGGCTTCCCAAT 5861  
QY 961 CCGGCGACAGTCCACATTTACTTGTGTCTTACCAAGCGCTCCAAATCATGTGGAAGCCGG 1020  
DB 5862 CCGGCGACAGTCCACATTTACTTGTGTGTCTTACCAAGCGCTCCAAATCATGTGGAAGCCGG 5921  
QY 1021 CAGTGCACGGAAGCTGTTCATCGGCTTACACAGGTTGCGGTGCTTGAAGGAC 1080  
DB 5922 CAGTGCACGGAAGCTGTTCATCGGCTTACACAGGTTGCGGTGCTTGAAGGAC 5981  
QY 1081 GGTTCATGAGCGAGATATGTGTGTGCTTCCGCTGTGCTATTAATCAAGCCCTTGA 1140  
DB 5982 GGTTCATGAGCGAGATATGTGTGTGCTTCCGCTGTGCTATTAATCAAGCCCTTGA 6041  
QY 1141 TCAGGACACATTTAGTGGGCTCATGACTCTGTGGCAAAATTTGGCTGAACGTTTGGGAG 1200  
DB 6042 TCAGGACACATTTAGTGGGCTCATGACTCTGTGGCAAAATTTGGCTGAACGTTTGGGAG 6101

QY 1201 GAGTCCTTCTTTCAGGATAGAGAGATCTTTCTGGGCAACATCTCTCGTGTGT -AA 1259  
DB 6102 GAGTCCTTCTTTCAGGATAGAGAGATCTTTCTGGGCAACATCTCTCGTGTGTGA 6161  
QY 1260 ACATGAGATTTTCTCTATGATTTGGGACCTTATCAAGCTTAATGGGAATAGATCTGGCGG 1319  
DB 6162 ACATGAGATTTTCTCTATGATTTGGGACCTTATCAAGCTTAATGGGAATAGATCTGGC -GG 6220  
QY 1320 GTTGGTCCAGTTTGTGAAGCGGGTTTATGAGATCTCTCCGTTGTGATCAACGGATA 1379  
DB 6221 GTTGGTCCAGTTTGTGAAGCGGGTTTATGAGATCTCTCCGTTGTGATCAACGGATA 6280  
QY 1380 TGAAGAAATCAACCGGATGTCCCTGGAAGAAATTCAGAACTTCCAGCTGTGATCCATC 1439  
DB 6281 TGAAGAAATCAACCGGATGTCCCTGGAAGAAATTCAGAACTTCCAGCTGTGATCCATC 6340  
QY 1440 TGAAGTGTAAACGCTGTCTGTGAGCCAGCGCAATATCCCATGTTCAAGTCAGGGCGAT 1499  
DB 6341 TGAAGTGTAAACGCTGTCTGTGAGCCAGCGCAATATCCCATGTTCAAGTCAGGGCGAT 6400  
QY 1500 TCAGAGGAAAGACAAATAATGAATGAAGCTTAAGACGGGATATCTGAATTTATGA 1559  
DB 6401 TCAGAGGAAAGACAAATAATGAATGAAGCTTAAGACGGGATATCTGAATTTATGA 6460  
QY 1560 TAAAGTGTGTCAATCTGCACTGCAAAATATCCAATCAGGCAATGCTGATGCA 1619  
DB 6461 TAAAGTGTGTCAATCTGCACTGCAAAATATCCAATCAGGCAATGCTGATGCA 6520  
QY 1620 TACCAATATTTTCAAGGACCACTGCAAAAGCGGTTTAAACAGCATATGACAGATC 1679  
DB 6521 TACCAATATTTTCAAGGACCACTGCAAAAGCGGTTTAAACAGCATATGACAGATC 6580  
QY 1680 GTCAAAACTCTTCTGATGATGATGAAGAAATTTCTGTTAGACATATCTCCGCTTTG 1739  
DB 6581 GTCAAAACTCTTCTGATGATGATGAAGAAATTTCTGTTAGACATATCTCCGCTTTG 6640  
QY 1740 TGTCTCATGAGAGGATGCAAAAGCAAGTATGCTGTGACTATGAGTGCAGATCC 1799  
DB 6641 TGTCTCATGAGAGGATGCAAAAGCAAGTATGCTGTGACTATGAGTGCAGATCC 6700  
QY 1800 GAATGTTAAAGTCTTACTGCTCATCACTTATATATGAGGAGAGACTCCCAAGCTGT 1859  
DB 6701 GAATGTTAAAGTCTTACTGCTCATCACTTATATATGAGGAGAGACTCCCAAGCTGT 6760  
QY 1860 GGCAGTCCCGCAAAAGAGGATTAATGTCTGTGCGGAGCGCAATTTGAGATCTTT 1919  
DB 6761 GGCAGTCCCGCAAAAGAGGATTAATGTCTGTGCGGAGCGCAATTTGAGATCTTT 6820  
QY 1920 CCGGCGGTTTGCACAGACCTATTTCTGTGCTGCGCTGATTAACGACCAAAATGTTATCA 1979  
DB 6821 CCGGCGGTTTGCACAGACCTATTTCTGTGCTGCGCTGATTAACGACCAAAATGTTATCA 6880  
QY 1980 ACATGATTTGGCTTACAGAGAAATGCGGGGAGCTTTCAACTCAACCGGGGTGTA 2039  
DB 6881 ACATGATTTGGCTTACAGAGAAATGCGGGGAGCTTTCAACTCAACCGGGGTGTA 6940  
QY 2040 GGAATTTTATTTGAGAAACTTTCTTTCAAGCACTGACACCGCTAATGATACCGGAGT 2099  
DB 6941 GGAATTTTATTTGAGAAACTTTCTTTCAAGCACTGACACCGCTAATGATACCGGAGT 7000  
QY 2100 TTAATTTGGCGGTTGCGAGTTTCTTCAAGGTGATGAGTGAAGGCTCTATTCAGAC 2159  
DB 7001 TTAATTTGGCGGTTGCGAGTTTCTTCAAGGTGATGAGTGAAGGCTCTATTCAGAC 7060  
QY 2160 CCGGTGTAACCGCTGTCTGTGCAATTAATCCCAATTTGTGAGGCAATTTTGGCAAGGCAA 2219  
DB 7061 CCGGTGTAACCGCTGTCTGTGCAATTAATCCCAATTTGTGAGGCAATTTTGGCAAGGCAA 7120  
QY 2220 TCCCTCGAACACTCTTGAAGAGATTAATCACTCCGAGTAGAATAATTAG 2268  
DB 7121 TCCCTCGAACACTCTTGAAGAGATTAATCACTCCGAGTAGAATAATTAG 7169

RESULT 6  
TIMPS2  
LOCUS  
DEFINITION  
T1 plasmid (A. tumefaciens octopine strain) tms1 gene, complete  
ACCESSION  
K02554  
VERSION  
K02554.1 GI:154747  
KEYWORDS  
2 of 2  
SEGMENT  
plasmid T1  
SOURCE  
plasmid T1  
REFERENCE  
1 (bases 1 to 2591)  
Klee, H., Montoya, A., Horodyski, F., Lichtenstein, C., Garfinkel, D.,  
Piller, S., Flores, C., Paschon, J., Nester, E. and Gordon, M.  
Nucleotide sequence of the tms genes of the pTiA6NC octopine Ti  
Proc. Natl. Acad. Sci. U.S.A. 81 (6), 1728-1732 (1984)  
JOURNAL  
MEDLINE  
PUBMED  
84170374  
COMMENT  
5584906  
Original source text: Plasmid T1 (clone: pTiA6NC).  
The tms1 protein has a region which showed significant homology to  
the adenine binding region of p-hydroxybenzoate hydroxylase from  
P. fluorescens. It seems that the tms1 protein binds adenine  
either as substrate or cofactor.  
The initiation and termination sites of transcription have not yet  
been identified [1].  
FEATURES  
source  
Location/Qualifiers  
1..2591  
/organism="Plasmid T1"  
/mol\_type="genomic DNA"  
/specific\_host="octopine"  
/db\_xref="taxon:2512"  
/clone="pTiA6NC."  
/plasmid="Plasmid T1"  
160..2427  
/gene="tms1"  
160..2427  
/gene="tms1"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AA92550.1"  
/db\_xref="GI:154750"  
/translation="MSASPLDNDCHLPKMYDLMVDKADLDRVSDAPLERAS  
RGRRTQISTECAGLACKRLADGRPEISAGKVAVLSYITIGKILGRILIESRPW  
ARATVGLVAIDLAPCMDFSEAOQLQALFLSGKRCAPLDISHFVAISIKYAGPRT  
LPMPLVENGTMKCVGTFTTLEGAVPDMVAGRNMLKGSASPTIDLLYRPF  
DCCSDGRIGFPEVDPKPKAVIAGISGLVAVNELNAGVDVTIYEASDVGKL  
WSHAFDADSVAVEMGMRFPAPAFCLPFELRYGLSSMRPFNPCTVDTIYVGGVO  
YMKAGQDLPKLEHRYNMGRAFLKDGFTFRDVLASPAVITDALKSGDIRAHDSMQ  
IMNRFGRSPFSSEIERIFLGHIPGGSTWSPFDWDLFKLMIGSGGFGVPESEFI  
ELIRLVINGEENORMCPBGISELPRRIASVENVGVSAQRICHVQVRAIQEKTKIK  
IRLKSGISLKYVVVTVSGLANIQLRHCLDINIPOAPVNOAVDSHMTGSKLPM  
TERKFWLDHILPSCVLMGLAKAVYCLDYEPDPPNGKGLVLIYTWEDSHKLLAVD  
KKERLLDLDAISRSFPAFOHLFPACADVDONVIOHDMLTDENAGAKRLNRGDF  
YSEELFFQALDPAISRTDNTGTVILAGSCSFTGGMWGAIQYACNAVCAIINHCGILANGN  
PLEHSWKRYNYRTRN"  
BASE COUNT  
ORIGIN  
654 a 570 c 656 g 711 t  
About 250 bp after segment 1.  
Query Match 98.1%; Score 2225.8; DB 1; Length 2591;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2255; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

Db 280 CGAAGAGCTTCTAAGGGAAGAGATTACTCAAAATCTCCACGAGTGCAGCGCTGGGTTA 339  
Qy 181 GCTTGCAGAAAGCTGGCCGATGTCCCTTCCCGAATCTCAGCTGTGAAAGGTAGCA 240  
Db 340 GCTTGCAGAAAGCTGGCCGATGTCCCTTCCCGAATCTCAGCTGTGAAAGGTAGCA 339  
Qy 241 GTTCTCCGCTATATCTATATATGGAAGAAATCTGAGGGGGGATCTGTAATGAAA 300  
Db 400 GTTCTCCGCTATATCTATATATGGAAGAAATCTGAGGGGGGATCTGTAATGAAA 459  
Qy 301 CTTGGGCGCGGCAACAGTGAAGTGTCTGTTGCCATCGACTTGACACCAATTTTCATG 360  
Db 460 CTTGGGCGCGGCAACAGTGAAGTGTCTGTTGCCATCGACTTGACACCAATTTTCATG 519  
Qy 361 GATTTCCTCGAAGACACAACTAATCCAGCCGTGTTTGTGTGACGGGTAAAGATGTCA 420  
Db 520 GATTTCCTCGAAGACACAACTAATCCAGCCGTGTTTGTGTGACGGGTAAAGATGTCA 579  
Qy 421 CGATTGATCTTATAGTATTTGCGGCAATTTTCAATCTTAAGCTGCGGCTTTCGAAC 480  
Db 580 CGATTGATCTTATAGTATTTGCGGCAATTTTCAATCTTAAGCTGCGGCTTTCGAAC 639  
Qy 481 CTGCCAATGCCGCTGTACAGAGATGAGACAGATGAATGCGTTTACCATTAAC 540  
Db 640 CTGCCAATGCCGCTGTACAGAGATGAGACAGATGAATGCGTTTACCATTAAC 699  
Qy 541 CTTGAAGGGGCGGTGCATTTGACATGTAGTATGTGTGAAACCTGATGTGAAGGT 600  
Db 700 CTTGAAGGGGCGGTGCATTTGACATGTAGTATGTGTGAAACCTGATGTGAAGGT 759  
Qy 601 TCGGCAAGTCTCTTCCAACTAATCGACTGTCTAGACTGACGACCGTTTTTGACCA 660  
Db 760 TCGGCAAGTCTCTTCCAACTAATCGACTGTCTAGACTGACGACCGTTTTTGACCA 819  
Qy 661 TGTTCGATAGTGAAGCAGATCCGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGTGGCG 720  
Db 820 TGTTCGATAGTGAAGCAGATCCGCTTCTTCCGAGAGATGTTCTTAAGCGAAAGTGGCG 879  
Qy 721 GTCAATGGCGCTGGCATTTCCGAGCTGTGTGAGCAACGAACTGTTATGTGGGTA 780  
Db 880 GTCAATGGCGCTGGCATTTCCGAGCTGTGTGAGCAACGAACTGTTATGTGGGTA 939  
Qy 781 GACGATTTAACAATATATGAAGCAAGTATCGTGTGAGGCAAGCTTTGTACATGCT 840  
Db 940 GACGATTTAACAATATATGAAGCAAGTATCGTGTGAGGCAAGCTTTGTACATGCT 999  
Qy 841 TTCAGGAGCGCTCTAGTGTGCGGCGGAAATGGGGGCGATGGATTTCTCTCGTGA 900  
Db 1000 TTCAGGAGCGCTCTAGTGTGCGGCGGAAATGGGGGCGATGGATTTCTCTCGTGA 1059  
Qy 901 TTCTGCTTTTCTTCTCTCGAGCGTTACGAGCTGTCTTCAATGAGCGCTTCCCAAT 960  
Db 1060 TTCTGCTTTTCTTCTCTCGAGCGTTACGAGCTGTCTTCAATGAGCGCTTCCCAAT 1119  
Qy 961 CCCGCGACGTGACACTTACTTGTGCTTACAGGCGTCCATATCATGTGAAAGCCGG 1020  
Db 1120 CCCGCGACGTGACACTTACTTGTGCTTACAGGCGTCCATATCATGTGAAAGCCGG 1179  
Qy 1021 CAGCTGCGACGAGGTGTTTCATCGCGTTTCAACGTTGAGGCGGCTTGAAGGAC 1080  
Db 1180 CAGCTGCGACGAGGTGTTTCATCGCGTTTCAACGTTGAGGCGGCTTGAAGGAC 1239  
Qy 1081 GGTTCATGAGCGAGATATGTGTGCTGCTGCTGCTATTAATCAGGCGCTTGAAA 1140  
Db 1240 GGTTCATGAGCGAGATATGTGTGCTGCTGCTGCTATTAATCAGGCGCTTGAAA 1299  
Qy 1141 TCAAGACATTAAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTCGGGAG 1200  
Db 1300 TCAAGACATTAAGTGGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTCGGGAG 1359  
Qy 1201 GAGTCCCTCTTCAAGGATAGAGAGATCTTTCGGGCAACATCTCTGCTGGTGT -AA 1259

|    |      |  |      |
|----|------|--|------|
| Db | 1360 | GAGTCTCTCTCTTCAAGGATAGAGAAGATCTTTCTGGGCAACAATCCTCCTGGTGTGGAA     | 1412 |
| Qy | 1260 | ACATGAGACTTTTCTCATGATTTGGGACCTATTCAAGCTAATGGAATNGATCTGGCGGG      | 1319 |
| Db | 1420 | ACATGAGATTTTCTCATGATTTGGGACCTATTCAAGCTAATGGAATNGATCTGGC -GG      | 1478 |
| Qy | 1320 | GTTTGGTCCATTTTTTGAAGCGGGTTTATTGAACTCTCCGCTTGGTCATCAACGGTA        | 1379 |
| Db | 1479 | GTTTGGTCCATTTTTTGAAGCGGGTTTATTGAACTCTCCGCTTGGTCATCAACGGTA        | 1538 |
| Qy | 1380 | TGAAGAAAATCAGCGGATGTGCCCTGAAGAAATCTCAGAACTTCCACGTCGATCGCATC      | 1439 |
| Db | 1539 | TGAAGAAAATCAGCGGATGTGCCCTGAAGAAATCTCAGAACTTCCACGTCGATCGCATC      | 1598 |
| Qy | 1440 | TGAAGTGTTAACGGTGTGTCTGTGAGCCAGCCGATATGCCATGTTCAAGTCAGGGCGAT      | 1499 |
| Db | 1599 | TGAAGTGTTAACGGTGTGTCTGTGAGCCAGCCGATATGCCATGTTCAAGTCAGGGCGAT      | 1658 |
| Qy | 1500 | TCAGAAAGAAAAGACAAAATTAAGAATAAGCTTAAGCCGGATATCTGAACTTTAAGA        | 1559 |
| Db | 1659 | TCAGAAAGAAAAGACAAAATTAAGAATAAGCTTAAGCCGGATATCTGAACTTTAAGA        | 1718 |
| Qy | 1560 | TAAAGTGTGTGCATCATGTGGACCTCGCAAAATATCCAATCTCAGGGCATTGCTCGACATGCGA | 1619 |
| Db | 1719 | TAAAGTGTGTGCATCATGTGGACCTCGCAAAATATCCAATCTCAGGGCATTGCTCGACATGCGA | 1778 |
| Qy | 1620 | TACCAATATTTTTTCAGGACCAAGTGAACCAAGCGGTTGATAACAGCCATATGACAGATC     | 1679 |
| Db | 1779 | TACCAATATTTTTTCAGGACCAAGTGAACCAAGCGGTTGATAACAGCCATATGACAGATC     | 1838 |
| Qy | 1680 | GTCAAAACCTCTTCTCGTGAATGTAACGAAATTTCTGGTTAACCATATCCTCCGCTCTTG     | 1739 |
| Db | 1839 | GTCAAAACCTCTTCTCGTGAATGTAACGAAATTTCTGGTTAACCATATCCTCCGCTCTTG     | 1898 |
| Qy | 1740 | TGTCCTCATGGACGGGATCGCAAAAGCAGTGAATTCCTGGACATATGAGTGGCAGATATCC    | 1799 |
| Db | 1899 | TGTCCTCATGGACGGGATCGCAAAAGCAGTGAATTCCTGGACATATGAGTGGCAGATATCC    | 1958 |
| Qy | 1800 | GAATGTTAAAGGTCTAGTGTCTCATTCAGTTATACATGGAGGACGACTCCCAAGCTGTT      | 1859 |
| Db | 1959 | GAATGTTAAAGGTCTAGTGTCTCATTCAGTTATACATGGAGGACGACTCCCAAGCTGTT      | 2018 |
| Qy | 1860 | GGCGGTCCCCGACAAAAGAGCCGATTATGTCTGCTCGGGACGCAATTTGGAATCTTT        | 1919 |
| Db | 2019 | GGCGGTCCCCGACAAAAGAGCCGATTATGTCTGCTCGGGACGCAATTTGGAATCTTT        | 2078 |
| Qy | 1920 | CCCCGCGTTTGGCCGACGACTATTTCCGCGCGCTGAGTATTAACGACCAAAATGTTATTCA    | 1979 |
| Db | 2079 | CCCCGCGTTTGGCCGACGACTATTTCCGCGCGCTGAGTATTAACGACCAAAATGTTATTCA    | 2138 |
| Qy | 1980 | ACATGATTGGCTTACAGACGAGAAATGCCGGGGAGCTTTCAAACCTCAACCGGCGTGTGA     | 2039 |
| Db | 2139 | ACATGATTGGCTTACAGACGAGAAATGCCGGGGAGCTTTCAAACCTCAACCGGCGTGTGA     | 2198 |
| Qy | 2040 | GGATTTTATTTCTGAAGAACTTTTCTTTCAGACACTGACACACGCTAATGATACCGGAGT     | 2099 |
| Db | 2199 | GGATTTTATTTCTGAAGAACTTTTCTTTCAGACACTGACACACGCTAATGATACCGGAGT     | 2258 |
| Qy | 2100 | TTACTTGGCGGGTTGCACTTGTCTTTCACAGGTGAATGGGTGGAGGGTGCATTTCAACAC     | 2159 |
| Db | 2259 | TTACTTGGCGGGTTGCACTTGTCTTTCACAGGTGAATGGGTGGAGGGTGCATTTCAACAC     | 2318 |
| Qy | 2160 | CGCGGTATACGCCGCTGTGTGCAAATATCCAAATTTGTGGAAGCATTTTGGCAAGGGCA      | 2219 |
| Db | 2319 | CGCGGTATACGCCGCTGTGTGCAAATATTCACAAATTTGTGGAAGCATTTTGGCAAGGGCA    | 2378 |
| Qy | 2220 | TCCTCTGCAACACTTCTGGAAGATATTAATTAACCGCAGTGAAGAAATTAG 2268         |      |
| Db | 2379 | TCCTCTGCAACACTTCTGGAAGATATTAATTAACCGCAGTGAAGAAATTAG 2427         |      |

|            |   |              |     |        |                 |
|------------|---|--------------|-----|--------|-----------------|
| LOCUS      | AE007927  | 9931 bp      | DNA | linear | BCT 18-DEC-2000 |
| DEFINITION | Agrobacterium tumefaciens str. C58 plasmid Ti, section 5 of 20 of the complete sequence.  |              |     |        |                 |
| ACCESSION  | AE007927  | AE007871     |     |        |                 |
| VERSION    | AE007927.2  | GI:156271969 |     |        |                 |
| KEYWORDS   |   |              |     |        |                 |
| SOURCE     | Agrobacterium tumefaciens str. C58 (Cereon)   |              |     |        |                 |
| ORGANISM   | Agrobacterium tumefaciens str. C58 (Cereon)   |              |     |        |                 |
| REFERENCE  | Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.   |              |     |        |                 |
| AUTHORS    | 1 (bases 1 to 9931)   |              |     |        |                 |
| TITLE      | Hinkle,G., Slater,S.C. and Goodner,B.   |              |     |        |                 |
| JOURNAL    | Complete Genome Sequence of Agrobacterium tumefaciens C58 (Rhizobium radiobacter C58), the Causative Agent of Crown Gall Disease in Plants  |              |     |        |                 |
| COMMENT    | Unpublished   |              |     |        |                 |
| FEATURES   | 2 (bases 1 to 9931)   |              |     |        |                 |
| source     | Hinkle,G., Slater,S.C. and Goodner,B.   |              |     |        |                 |
|            | Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street, Cambridge, MA 02133, USA   |              |     |        |                 |
|            | On Oct 19, 2001 this sequence version replaced gi:15163470.   |              |     |        |                 |
|            | location/Qualifiers   |              |     |        |                 |
|            | 1. 9931   |              |     |        |                 |
|            | /organism="Agrobacterium tumefaciens str. C58 (Cereon)"   |              |     |        |                 |
|            | /mol_type="genomic DNA"   |              |     |        |                 |
|            | /strain="C58"   |              |     |        |                 |
|            | /db_xref="taxon:181661"   |              |     |        |                 |
|            | /plasmid="Ti"   |              |     |        |                 |
|            | 240. .905   |              |     |        |                 |
|            | /gene="AGR_pti_bx83"  |              |     |        |                 |
|            | 240. .905   |              |     |        |                 |
|            | /gene="AGR_pti_bx83"  |              |     |        |                 |
|            | /note="putative indole-3-lactate synthase"  |              |     |        |                 |
|            | /codon_start=1  |              |     |        |                 |
|            | /transl_table=1   |              |     |        |                 |
|            | /product="AGR_pti_bx83p"  |              |     |        |                 |
|            | /protein_id="AAK0967.1"   |              |     |        |                 |
|            | /db_xref="GI:15163471"  |              |     |        |                 |
|            | /translation="MYHSRPIFINIDSNNIOBRRELKLVLRHTEIAYRSPAOEDLIPARORSMSIINTDVPIDPAIDEVYRCEVACLPGPACIPLNIILNDELTVYVCSFQAMRKVAHRKFDVGSDEGVISTVPEYAGIKRETRMSNNVCONTSETHDLDAIYAPLRPSLQNPSPSHKIKGCDSPFLASRPDPFCEVETIAYVKALFHNGLKKNPRCGGQMLP   |              |     |        |                 |
|            | PSC"  |              |     |        |                 |
|            | complement (1410. .2813)  |              |     |        |                 |
|            | /gene="AGR_pti_48"  |              |     |        |                 |
|            | complement (1410. .2813)  |              |     |        |                 |
|            | /gene="AGR_pti_48"  |              |     |        |                 |
|            | /note="(AJ237588) IMH"  |              |     |        |                 |
|            | /codon_start=1  |              |     |        |                 |
|            | /transl_table=1   |              |     |        |                 |
|            | /product="AGR_pti_48p"  |              |     |        |                 |
|            | /protein_id="AAK0968.1"   |              |     |        |                 |
|            | /db_xref="GI:15163472"  |              |     |        |                 |
|            | /translation="MVPITSLAQTLELRKRDYSCLEIVETLIARCOAKPKNALLATDWDQLRSRAKIDHGNAGLGCIPICFPAANIATGIFPISATPALINHDPKIPSVWAEERFISAGALPGASGNMHELSFGITSNNTVAIGAVRNWNPSSLPGSGSVAAVASRLMIGGIGTGTGASVRLPALCGVGFPTLARYPRDRIIIVSPRTDAGIIAOCVADVIVIIIDQVIGSRGSAKISPMWILGRLGIPPTVEYDLDADVAFAATTIRLANRGVTEAEADIPPIBELNSGASIPALVEFPHALKGLIDPVGCVSPSDVTKGIRSPDANITVAQIDHQIISNDEYELAROSFRRLDATTNRNFRLYQDALILFPTPLAAKIGGSSVYIHNGSMWNTFKIYYANVDPSSNAGLPGSLPACLPRLPVGMETIDGLAGDHRLLAIGAALEKALNPSPFPAFN" |              |     |        |                 |
|            | 3151. .5418   |              |     |        |                 |
|            | /gene="AGR_pti_49"  |              |     |        |                 |
|            | 3151. .5418   |              |     |        |                 |
|            | /gene="AGR_pti_49"  |              |     |        |                 |
|            | /note="(AF126446) tryptophan monooxygenase"   |              |     |        |                 |
|            | /codon_start=1  |              |     |        |                 |
|            | /transl_table=1   |              |     |        |                 |
|            | /product="AGR_pti_49p"  |              |     |        |                 |
|            | /protein_id="AAK0969.1"   |              |     |        |                 |
|            | /db_xref="GI:15163473"  |              |     |        |                 |

/translation="MSASALLDNCDFSTKMDLINVDADEHREVSDFSERAS  
GRRITQISGSCAGLACKLADRPPEISAGEKVALASVYVKGELIRLIESEBM  
ARARVGLVAILDLAPFCMDPSEADQLQTLFLSGKRASSDLSPFVALISISKARSP  
LQMPYKCTTKTKVTGFTLTLIEAVFPDMAYGNMLKKSAGSFPTIDLLYTRSF  
DQCSGRIGFPEIDVPRKVALIAGISGLVASSELHAGVDVITLYEASDRVGKL  
WSHAFKAPSVVAMGMRPPASCFEFLERGLASMRPPMPGTVDNLYOGILR  
YMKAGQOPKLFHRYVSGWRALFKDFHGDIVLAPVALIOLAKSGDIRRAHDSQ  
TYMNRFGRESFSAIERIFLGTHPGGETWSFPDMDLFLMIGISGFGPFPESGPI  
ELLRLVINGEENORMCSGISELPRILAIQVYVNSVSGORISHVQRAIEKERTIK  
IRLKSGLSELYDKVYVTSGLANTOLRHCTCDTTIPAPVQAVIDNSHMTGSLFL  
TERFVLDHILPSCVLMGDIKAVYCLDYEPPQDNGGLVLIISYWEDESHKLAVD  
KGRFCLRLDAISRFPALVACADYDONVQHDWLTDENAAGAFPLNREGEF  
YSELFEQALDMTQDGVLYLAGSCSFSGMWEGAIOACNAVCAIHNCGILLAKON  
PLEHSMKRYVYRNRN"  
6089 .6838  
/gene="AGR\_pt1\_50"  
6089 .6838  
/gene="AGR\_pt1\_50"  
/note="Probable tmr gene"  
/codon\_start=1  
/product="AGR\_pt1\_50p"  
/protein\_id="AAK90970.1"  
/db\_xref="GI:15163474"  
/translation="MYSRYRKLMIDLRIEPTCTGKTSYAVLAQOTGLPVLSDRV  
QCCPOLSTGSGRPTVEELKGTSLRYLDRLPLVGLIAKQAHRIEMGEVYVNAHGL  
ILEGGSILKCMQASVMSADPRWHIIRLELADERTHNVAVAKRYQMKLRPAAGSI  
IOELVDLMEKPRILPILKEIDGTRYAMLFPASQNTSDMLLDLADMEDLHIGLAE  
YLIHARRQEQKFPVNAAYDGFGEHFGMY"  
7354 .7914  
/gene="AGR\_pt1\_51"  
7354 .7914  
/gene="AGR\_pt1\_51"  
/note="Probable 6a gene"  
/codon\_start=1  
/product="AGR\_pt1\_51p"  
/protein\_id="AAK90971.1"  
/db\_xref="GI:15163475"  
/translation="MTRQGMVLPCLSHKDDQLODELSLSKAYRRKKFQDTLHTKSGD  
IIDPSGEFLVYLDENYRICTQSMVLSVAPDGLATTLEPVSQSYTFQVBOLOA  
FGSDGRINYSRNEHSSYFLTIQASNEFHVGA VNTFGSKDIWRKMPDASQPLD  
YHLIIVGCSAFLEBALDVEDSDAV"  
complement (8397..9020)  
/gene="AGR\_pt1\_bx9p"  
complement (8397..9020)  
/gene="AGR\_pt1\_bx9p"  
/note="tm1"  
/codon\_start=1  
/product="AGR\_pt1\_bx9p"  
/protein\_id="AAK90972.1"  
/db\_xref="GI:15163476"  
/translation="MTVANQVRLDTLILRTGEMOSRLBOARTFGALLPEYVFOFS  
AIRLGEDEYILTRQELVYVLRREDIARCALRHLRHPSSNSGIMATYIPWLMDA  
RLINRYMOERCDOGLHYHQGBHTNOFLAIMPNSCFAPRSDVYNENRYGYANGG  
NYTEBGEDEDEDEBDEBTGAEIRDSQTNLINTPIALGSCHLSA"  
BASE COUNT 2639 a 2177 c 2364 g 2751 t  
ORIGIN  
Query Match 90.1% Score 2044; DB 1; Length 9931;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 2141; Conservative 0; Mismatches 125; Indels 2; Gaps 2;

Db 3271 CGGAAGCTTCTAGGGGAAGAGATTACTCAATCTCCGGCAGAGTGCACGCTGGTTA 3330  
Qy 181 GCTTGCAAAAAGCTGGCCGATGTCCTTCCCGAGATCTCACCTGGTGAAGTAGCA 240  
Db 3331 GCTTGCAAAAAGCTGGCCGATGTCCTTCCCGAGATCTCACCTGGTGAAGTAGCA 3390  
Qy 241 GTTCTCCGCTTATATCTATATGGCAAGAAATCTGGGGGGGATCTTGAATCGAA 300  
Db 3391 GCCCTCTCCGCTTACATATATGTTGGCAAGAAATCTGGGGGGGATCTTGAATCGAA 3450  
Qy 301 CTTTGGGCGGGGCAACAGTAGTGTCTGTTGCCATCGACTTGGACACCATTTGGATG 360  
Db 3451 CTTTGGGCGGGGCAACAGTAGTGTCTGTTGCCATCGACTTGGACACCATTTGGATG 3510  
Qy 361 GATTTCTCCGAGACACAACTATCCAAAGCCGTGTTTTGCTGAGCGGTAAAGATGCA 420  
Db 3511 GATTTCTCCGAGACACAACTATCCAAAGCCGTGTTTTGCTGAGCGGTAAAGATGCA 3570  
Qy 421 CCGATTGATCTTATGATTTGCGTGGCCATTTGCAATCTTAAAGCTCGCGCTTCCAAAC 480  
Db 3571 TCCAGGATCTTATGATATTCGTCGTCATTTCAATCTTAAAGCTCGCGCTTCCAAAC 3630  
Qy 481 CTGCCAATGCGCGCTGACGAGAAATGCGACGATGAATGCGTTACCGGGTTTACATTAAC 540  
Db 3631 CTGCCAATGCGCGCTTACGAGAAAGCAGACGAAACGCGTTACCGGGTTTACCTGACC 3690  
Qy 541 CTTGAAAGGGGCGTGCATTTGACATGATGATGTTAGTGTGAAACCTGATGCTGAAGGT 600  
Db 3691 CTTGAAAGGGGCGTGCATTTGACATGATGATGTTAGTGTGAAACCTGATGCTGAAGGT 3750  
Qy 601 TCGGCAAGTTCTCTTCCAAACATCGACTTGTCTAGACTGCGAGACCGTTTTTGACCA 660  
Db 3751 TCGGCAAGTTCTCTTCCAAACATCGACTTGTCTAGACTGCGAGACCGTTTTTGACCA 3810  
Qy 661 TGTTCGATAGTGAAGCATTCGCTCTTCCGGAAGATGTTCTTAAGCCGAAGTGGCG 720  
Db 3811 TGTTCGATAGTGAAGCATTCGCTCTTCCGGAAGATGTTCTTAAGCCGAAGTGGCG 3870  
Qy 721 GTTCATTGGCGCTGGCATTTCCGCACTGTGTGGCAACGAACTGCTTCACTGGGGTA 780  
Db 3871 ATCATTGGCGCTGGCATTTCCGCACTGTGTGGCAACGAACTGCTTCACTGGGGTA 3930  
Qy 781 GACGATGTTACATATATGAAGCATGATCGTGTGGAGCAAGCTTTGGTCAATGCT 840  
Db 3931 GACGATGTTACATATATGAAGCATGATCGTGTGGAGCAAGCTTTGGTCAATGCT 3990  
Qy 841 TTCAAGGACGCTCTGTGTGCGGCGCAAAATGGGGGCGATGCGATTTCTCTGCTGCA 900  
Db 3991 TTCAAGGACGCTCTGTGTGCGGCGCAAAATGGGGGCGATGCGATTTCTCTGCTGCA 4050  
Qy 901 TTCTGCTTGTGTTTCTTCTCGAGCGTTACGCGCTGTCTTCAATGAGCGCTTCCCAAT 960  
Db 4051 TTCTGCTTGTGTTTCTTCTCGAGCGTTACGCGCTGTCTTCAATGAGCGCTTCCCAAT 4110  
Qy 961 CCGGCGACGTCGACACTTACTTGTCTTCAAGCGCTTCCATATCATGTGGAAGCCGG 1020  
Db 4111 CCGGCGACGTCGACACTTACTTGTCTTCAAGCGCTTCCATATCATGTGGAAGCCGG 4170  
Qy 1021 CAGCTGCAACGAGCTTTCATCGCGTTTCAAGCGTTTGGGCGGCTTCTTGAAGGAC 1080  
Db 4171 CAGCTGCAACGAGCTTTCATCGCGTTTTCATCGCGTTTTCATCGCGTTTTCATCGCG 4230  
Qy 1081 GGTTCATGAGCGAGATATGTTGAGCTTGGCTGTGCGCTTATTAAGCGCTTGA 1140  
Db 4231 GGTTCATGAGCGAGATATGTTGAGCTTGGCTGTGCGCTTATTAAGCGCTTGA 4290  
Qy 1141 TCAAGCAATTAAGTGGGCTCATGACTCTGGCAAAATTTGGTGAACGTTTGGGAGG 1200  
Db 4291 TCAAGCAATTAAGTGGGCTCATGACTCTGGCAAAATTTGGTGAACGTTTGGGAGG 4350  
Qy 1201 GAGTCTTCTTCAAGGATAGAGAGATCTTTCGGGACACATCTCTGGGTGT -AA 1259

Db 4351 GAGTCTCTCTTCTTACGCGATAGAGAGATCTTTCTGGGACGCACTCCCTGGTGATGA 4410  
 QY 1260 ACATGAGTTTCTCTATGATTTGGGACCTATTCAAGCTATGGGAATGAGATCTGGCGG 1319  
 Db 4411 ACATGAGTTTCTCTATGATTTGGGACCTATTCAAGCTATGGGAATGAGATCTGGC-GG 4469  
 QY 1320 GTTTGGTCCAGTTTGGAAAGCGGGTTTATGAGATCTCTCGGTGGTGCATCAACGATTA 1379  
 Db 4470 GTTTGGTCCAGTTTGGAAAGCGGGTTTATGAGATCTCTCGGTGGTGCATCAACGATTA 4529  
 QY 1380 TGAAGAAATACAGCGGATGCTCTGAAGGAATCTCAGAACTTCCAGCTGCATCCGATC 1439  
 Db 4530 TGAAGAAATACAGCGGATGCTCTGAAGGAATCTCAGAACTTCCAGCTGCATCCGATC 4589  
 QY 1440 TGAAGGTTTAAAGGTTGCTGTGAGCCGCGCATATGCGATTCAGTCCAGTGGCGAT 1499  
 Db 4590 TCAAGGTTTAAAGGTTGCTGTGAGCCGCGCATATGCGATTCAGTCCAGTGGCGAT 4649  
 QY 1500 TCAGAGAAAGAAAGACAAAATAAAGATTAAGGCTTAAGAGCGGGATATCTGAATTTATGA 1559  
 Db 4650 TGAAGAAAGAAAGACAAAATAAAGATTAAGGCTTAAGAGCGGGATATCTGAATTTATGA 4709  
 QY 1560 TAAAGTGTTGTCATCTGAGACTCGCAATATCCAACTCAGGCAATTGCTGATCCGA 1619  
 Db 4710 TAAAGTGTTGTCATCTGAGACTCGCAATATCCAACTCAGGCAATTGCTGATCCGA 4769  
 QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGGGTTGTAAGACCATATGACAGATC 1679  
 Db 4770 TACCAATATTTTTCAGGACCAAGTGAACCAAGGGTTGTAAGACCATATGACAGATC 4829  
 QY 1680 GTCAAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1739  
 Db 4830 GTCAAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4889  
 QY 1740 TGTCTCATGAGACGGATGCGCAAAAGCATGATGCTGAGTATGATGATGATGATGATG 1799  
 Db 4890 TGTCTCATGAGACGGATGCGCAAAAGCATGATGCTGAGTATGATGATGATGATGATG 4949  
 QY 1800 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1859  
 Db 4950 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5009  
 QY 1860 GGGGCTCCCGACAAAAGAGGATTAATGCTGCTGCGGACGCAATTCGAGATCTTT 1919  
 Db 5010 GGGGCTCCCGACAAAAGAGGATTAATGCTGCTGCGGACGCAATTCGAGATCTTT 5069  
 QY 1920 CCGGCGTTTGGCCGACCACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979  
 Db 5070 CCGGCGTTTGGCCGACCACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5129  
 QY 1980 ACATGATTTGGCTTACAGACGAAATGCGGGGAGCTTCAAACTCAACCGGCGTGTGA 2039  
 Db 5130 ACATGATTTGGCTTACAGACGAAATGCGGGGAGCTTCAAACTCAACCGGCGTGTGA 5189  
 QY 2040 GGAATTTTATTTGAGAACTTTCTTTCAGACATGACACGCGTAAATGATACCGGAGT 2099  
 Db 5190 GGAATTTTATTTGAGAACTTTCTTTCAGACGCGTAAATGATACCGGAGT 5249  
 QY 2100 TTAATTTGCGGGTGGAGTTTCTTTCAGACGATGATGATGATGATGATGATGATGATG 2159  
 Db 5250 TTAATTTGCGGGTGGAGTTTCTTTCAGACGATGATGATGATGATGATGATGATGATG 5309  
 QY 2160 CCGGTTTAAACCGCTGCTGAGCAATTTCCCAATTTGAGGATTTGGCAAGGCA 2219  
 Db 5310 CCGGTTTAAACCGCTGCTGAGCAATTTCCCAATTTGAGGATTTGGCAAGGCA 5369  
 QY 2220 TCTCTCGAAGACTTTGGAAGATATACTACCGCAGTAAATTA 2267  
 Db 5370 TCTCTCGAAGACTTTGGAAGATATACTACCGCAGTAAATTA 5417

RESULT 8  
 AB009419

LOCUS AB009419 10377 bp DNA linear BCT 20-DEC-2001  
 DEFINITION Agrobacterium tumefaciens str. C58 Ti plasmid, section 2 of 21 of the complete sequence.  
 ACCESSION AB009419 AB008690  
 VERSION AB009419.1 GI:17744019  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Agrobacterium tumefaciens str. C58 (U. Washington)  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 1 (bases 1 to 10377)  
 Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Zhou, Y., Woo, L., Kiteajima, J.P., Okura, V.K., Almeida Jr., N.F., Gillet, M., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Gordon, D., Eissen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.  
 Direct Submission  
 Submitted (27-SEP-2001) Department of Microbiology, University of Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA 98195-7242, USA  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 11743193  
 2 (bases 1 to 10377)  
 Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E., Chen, Y., Woo, L., Kiteajima, J.P., Okura, V.K., Almeida Jr., N.F., Zhou, Y., Boyee Sr., D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, M., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Gordon, D., Eissen, J.A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V. and Nester, E.W.  
 The genome of the natural genetic engineer Agrobacterium tumefaciens C58  
 Science 294 (5550), 2317-2323 (2001)  
 21608550  
 11743193  
 REFERENCE  
 AUTHORS  
 FEATURES  
 source  
 Location/Qualifiers  
 1..10377  
 /organism="Agrobacterium tumefaciens str. C58 (U. Washington)"  
 /mol\_type="genomic DNA"  
 /strain="C58"  
 /db\_xref="taxon:180835"  
 /plasmid="T1"  
 319..513  
 /gene="ipt"  
 /note="synonym: Atu6006"  
 319..513  
 /gene="ipt"  
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"  
 /codon\_start=1  
 /transl\_table=11  
 /product="isopentenyladenosine phosphotransferase"  
 /protein\_id="AA146242.1"  
 /db\_xref="GI:17744020"  
 /translation="MDLYLIFGPTCTGKTAATLSEHTGIPVLAIDRIQCPLSTGSGRPAAYTAPQSIGORC"  
 complement(380..1414)  
 /gene="mas1"  
 /note="synonym: Atu6007"  
 complement(380..1414)  
 /gene="mas1"  
 /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer"  
 /codon\_start=1  
 /transl\_table=11  
 /product="agrocinnopine synthesis reductase"

/protein\_id="AAL46243.1"  
/db\_xref="GI:17744021"  
/translation="MEGGGLPRMFVENVPCDEKTELFSARVAGALTARRSNIILEV  
SHGVLSVATLIGVELTAMORCNACVILHFKPGSSMNVTDRSGVILVYAGSGIDK  
AIDNIMNRGRFSLGARSVEKLEATGFQNESLHRRFAEDLOTEEVVSTAMIK  
GRIDLVNNGDGRVALAAGIDIEQLERWKSNAGAPLIMLCLEPOLKSSGRIIV  
NIISMGQVRVLSFVGNMTKALSGTLTKAHOVWGESGVRTWEICPGVATQMSMT  
DEIDEGKIQPEHIAELVROAIERPNNAFVPIKEVLICIQPLAVRSRSKVGSIYVGP  
KPRGCCACAA"  
/complement(1622..2878)  
/gene="acscx"  
/note="synonym: Atu6008"  
/complement(1622..2878)  
/gene="acscx"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=11  
/product="agrocinnopine synthase"  
/protein\_id="AAL46244.1"  
/db\_xref="GI:17744022"  
/translation="MEGTWPMTLHMDLPENYTPPEVLAHNSVPRLEKLEEDLPQ  
VIRGMFNLGKRIQECTASSLALAGQGRNISELDVCLTSDNVAIVSHDLNTWVS  
EKLGDKLFNEHSSKIKDVPVIRVNSNGIIOKYLETIDHILPLEIFSKPLANS  
ATIFLQGRNEAIVAMLSHREYQVRVVLFTYEYPHGAFVDAVLAQASAMR  
KSIJALMPALPEELCLARLOYTEPTVDLILAGKAMPMSLMODMRIAHAHVSG  
VFNILIGOVVDKVLAFDSDOAVALAYLYKEDTMRARPHLKRAVTRCYDPAL  
LDSEGERSEIDIKTGRARHETDERKHIMKRGTSNGATIDWVTSIDREDEMAIW  
EMNQGIDREVSLSPLDLNLTISK"  
/complement(3636..4157)  
/gene="genes5"  
/note="synonym: Atu6009"  
/complement(3636..4157)  
/gene="genes5"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=11  
/product="gene 5 protein"  
/protein\_id="AAL46245.1"  
/db\_xref="GI:17744023"  
/translation="MNSIINTDVPIDPAIDEVKRCVACLPAGAPIPLIILNDLS  
TYYVCSFOAMRKAKHAKRFYDVGDEGVISTVPVAGITKEMRSMHNVCONTSNE  
THLDVIAVLFPTSLQNPFSHMKIGODSLARSDPFCVELIANGKALFHNGLAK  
NPRCGQWLPBSC"  
/complement(4662..6065)  
/gene="tms2"  
/note="synonym: Atu6010"  
/complement(4662..6065)  
/gene="tms2"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=11  
/product="indole acetamide hydrolase"  
/protein\_id="AAL46246.1"  
/db\_xref="GI:17744024"  
/translation="WVPTSLAQTLERLRKDYSCLELVETLIARCOAKPLNALIAT  
DMWGLRSAAKIDRHNGAGLGLGIPLCFANATGIPPSAATPALIHLPITPSRV  
AERLPSAGALPGASGMHLEISFGITSNNTAVNAPNPSLIPGSSGVAIAVSR  
LMIGIGTGTGASVRLPALCGVVGFRPTARVPRDIIPVSPRTAGIAGCVADV  
IILDVYISGSAKISPMPLKGLRIGLPTTYFVDDLDADVAFAETIRILLANGVTV  
EADIPLHELSNGASLPALYERPHALKITLDDFVGTGSFSDVIKIRSDVANIYSA  
QIDGHOISDEYELAROSFRPLQATYRNFRVYOLDALIFPAPLAARAIGQESSVI  
HNGSMNTEKTIYRVNDPSSNAGLPGLSLPACLTLPRLPVGMEIDGLASDRLLAIG  
AALAKAINPSPFDATN"  
/complement(6403..8670)  
/gene="tms1"  
/note="synonym: Atu6011"  
/complement(6403..8670)  
/gene="tms1"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"

/codon\_start=1  
/transl\_table=11  
/product="tryptophan 2-monooxygenase"  
/protein\_id="AAL46247.1"  
/db\_xref="GI:17744025"  
/translation="MSASALDNOQDFSTKRVLDIMVDAKDELDRVSDAFSEBEAS  
RGRITQIIGESGAGLACRLADGRPEISAGKRVAAIYVGEIIGRIILIESPW  
ARRAVGALVADLAPPCMSSEAKLOTLTFLSGKRCASDLSHPVAISIKSARSPT  
LQMPREKCTTKRVGFTLLEAVPDMVAAGRNIMLAKASGSPFTILLVYRSFP  
DQCSBSGRIGFPPEVDYPRKVAIIGAGISGLVYASLELHAGCDVTTIYASDRVCGKL  
WHAERDAPSVVAEWGAMRFPPASCLFEFLERYGLSSMRPEPNPTVTNLYOGLR  
YMKAGQOPPLKFRVYSGMRAFLKXDFHEGDIVLASPYALTOALSGDIIRAHDSMO  
TWNLFGRSFSFSAIERIFLGTHPGSGTSEPHMDLFLKMGISGSGPVPESGFI  
ELRLVINYENORMSSEGISLPRIRATOVNVGVSORIRHVORVRAIEKTKIK  
ILKSGISLKYKVVYVTSGLANTOLPHCTTCOTTPRAVNOAVDMSHTGSKLFL  
TERKRWLHILPSCVLMDGIKAVYCLDPEPODPNGKGLVLISTYEDSHKULAVPD  
KKEPFCLRLDAISRSFPAPAOHLVPAQCAVDONVVOHDLTBNAGAFKLNRRGDF  
YSEELFFQALDNTDITGVYLAQSCSFTGQWEGAIQTCANVCAIIRHNCGLIAXDN  
PLEHSWKRYNRYNRN"  
/complement(9368..10090)  
/gene="ipc"  
/note="synonym: Atu6012"  
/complement(9368..10090)  
/gene="ipc"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer"  
/codon\_start=1  
/transl\_table=11  
/product="isopentenyl transferase"  
/protein\_id="AAL46248.1"  
/db\_xref="GI:17744026"  
/translation="MDLRLIGPTCTGKTSTAVALAONOTGLPVLISDRVOCROSTG  
SGRPYBEHAKTSRLYLDPRPYKGIILAKQKQHERLMGEVYVYEAHAGILBEGSISL  
LKQMOSSYWSADFRRHITRHELDAPETMNAKARVQMLPRAGLSTIIOELVDLWK  
BEPLRLIKEIDGYRYAMLFAQONQITSDMLQLDADMDKLIHGAOBYLIHAROE  
QKPPVNAAYVGFEGHPGMY"  
/complement(2376..2438)  
BASE COUNT 2745 a 2376 c 2438 g 2818 t  
ORIGIN  
Query Match 90.1%; Score 2044; DB 1; Length 10377;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 2141; Conservative 0; Mismatches 125; Indels 2; Gaps 2;  
QY 1 ATGTGAGTTCACTCTCTGATATACGATGATCACTCCCAACCAAAATGRTGAT 60  
DB 6403 ATGTGAGTTCACTCTCTGATATACGATGATCACTCCCAACCAAAATGRTGAT 6462  
QY 61 CTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
DB 6463 CTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6522  
QY 121 CGAAGAGCTTTAGAGGAGAGAGATTAACAAGCTCCACGAGTGCAGCGCTGGGTTA 180  
DB 6523 CGAAGAGCTTTAGAGGAGAGAGATTAACAAGCTCCGAGAGTGCAGCGCTGGGTTA 6582  
QY 181 GCTTGAAGAGAGCTGGCCATGCTGCTTCCCGAGATCTCAGCTGTGAAAAGGTAGCA 240  
DB 6583 GCTTGAAGAGAGCTGGCCATGCTGCTTCCCGAGATCTCAGCTGTGAAAAGGTAGCA 6642  
QY 241 GTTCTCTCGCTTATATATATATGGAAGAAATCTGGGCGGATATCTTGAATGAAA 300  
DB 6643 GCCCTCTCGCTTATATATATGGAAGAAATCTGGGCGGATATCTTGAATGAAA 6702  
QY 301 CTTTGGCGCGGAGCAACAGTAGTGTCTGTTGCCATGACCTTGACACCAATTTGCATG 360  
DB 6703 CTTTGGCGCGGAGCAACAGTAGTGTCTGTTGCCATGACCTTGACACCAATTTGCATG 6762  
QY 361 GATTTCTCGAGACCAATATCCAGGCGCTTTTCTGAGCGGTAAAGATGTGCA 420  
DB 6763 GATTTCTCGAGACCAATATCCAGGCGCTTTTCTGAGCGGTAAAGATGTGCA 6822  
QY 421 CGAATTGATCTAGCATTTCTGGGCACTTCAATCTCAAACTGCGCGCTTTGCAAC 480

Db 6823 TCCAGCGATCTTATGATCTTGTGGCCATTTCAATCTTAAGACGTGCCCGCTCCGAACC 6882  
 Qy 481 CTGCAATGCGCGCTGTACGAGATGAGACGATGAAAATGCGTTACCGGGTTTACCATTAAC 540  
 Db 6883 CTGCAATGCGCGCTGTACGAGATGAGACGATGAAAATGCGTTACCGGGTTTACCATTAAC 6942  
 Qy 541 CTGGAAGGGGCGTGGCCATTTGACATGATGATTAATGTTGAAAACCTGATGCTGAAGGT 600  
 Db 6943 CTGGAAGGGGCGTGGCCATTTGACATGATGATTAATGTTGAAAACCTGATGCTGAAGGT 7002  
 Qy 601 TCGGCGAGTTCCTTTCCAAACATCGACTTGTCTAAGACCTGCAAGCCGTTTTTTGACCA 660  
 Db 7003 TCGGCGAGTTCCTTTCCAAACATCGACTTGTCTAAGACCTGCAAGCCGTTTTTTGACCA 7062  
 Qy 661 TGTTCGATGTGAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720  
 Db 7063 TGTTCGATGTGAGCGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 7122  
 Qy 721 GTCAATGGCGCTGGCAATTTCCGACCTGTTGGTGGCAACGAACTGCTTCATGCTGGGGTA 780  
 Db 7123 ATCAATGGCGCTGGCAATTTCCGACCTGTTGGTGGCAACGAACTGCTTCATGCTGGGGTA 7182  
 Qy 781 GAGCATGTTTCAATATATGAGACGATGATCGTTGTGAGGCAAGCTTTGCTCAATGCT 840  
 Db 7183 GAGCATGTTTCAATATATGAGACGATGATCGTTGTGAGGCAAGCTTTGCTCAATGCT 7242  
 Qy 841 TTCAAGGAGCGCTCTAGTGTGTGGCGGAAATGGGGGCGATGCGATTTCTCTCGCTGCA 900  
 Db 7243 TTCAAGGAGCGCTCTAGTGTGTGGCGGAAATGGGGGCGATGCGATTTCTCTCGCTGCA 7302  
 Qy 901 TTCTGCTTTGTTTCTTCTCGACCGTTTACGCGCTGTCTTTCATGAGGCGGTTCCCAAT 960  
 Db 7303 TCTGCTGTTTCTTCTTCTCGACCGTTTACGCGCTGTCTTTCATGAGGCGGTTCCCAAT 7362  
 Qy 961 CCGGCGCAAGTCCGACCTTCTGTTTACCAAGGCGTCCAAATCATGTTGAAAAGCGGG 1020  
 Db 7363 CCGGCGCAAGTCCGACCTTCTGTTTACCAAGGCGTCCAAATCATGTTGAAAAGCGGG 7422  
 Qy 1021 CAGCTGCGACGAGCTGTTCCATCGCGTTTACCAAGGCGTGGCGGTTCTTGAAGGAC 1080  
 Db 7423 CAGCTGCGACGAGCTGTTCCATCGCGTTTACCAAGGCGTGGCGGTTCTTGAAGGAC 7482  
 Qy 1081 GGTTCATGAGCGAGATATGTTGTTGGCTTCCGCTGTGCTATTAATCAAGCCCTTGA 1140  
 Db 7483 GGTTCATGAGCGAGATATGTTGTTGGCTTCCGCTGTGCTATTAATCAAGCCCTTGA 7542  
 Qy 1141 TCAGGACATTAAGTGGGCTCATGACTCTCGGCAATTTGGCTGAACCGTTTGGGAGG 1200  
 Db 7543 TCAGGACATTAAGTGGGCTCATGACTCTCGGCAATTTGGCTGAACCGTTTGGGAGG 7602  
 Qy 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGGGCAACATCTCTCGGAGTGT-AA 1259  
 Db 7603 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGGGCAACATCTCTCGGAGTGT-AA 7662  
 Qy 1260 ACATGAGTTTCTCATGATTTGGGACCTTATTCAGCTTAATGGAGATAGATCTGGCGG 1319  
 Db 7663 ACATGAGTTTCTCATGATTTGGGACCTTATTCAGCTTAATGGAGATAGATCTGGCGG 7721  
 Qy 1320 GTTTCGTCAGTTTGAAGGCGGTTTATGAGATCTCTCGGCTTGGTCAATCAAGGATA 1379  
 Db 7722 GTTTCGTCAGTTTGAAGGCGGTTTATGAGATCTCTCGGCTTGGTCAATCAAGGATA 7781  
 Qy 1380 TGAAGAAATCAAGGATAGAGGATCTGAGAAATCTCAAGCTTCAAGTGGATCGGATC 1439  
 Db 7782 TGAAGAAATCAAGGATAGAGGATCTGAGAAATCTCAAGCTTCAAGTGGATCGGATC 7841  
 Qy 1440 TGAAGTGTAAAGGATAGAGGATCTGAGAAATCTCAAGCTTCAAGTGGATCGGATC 1499  
 Db 7842 TGAAGTGTAAAGGATAGAGGATCTGAGAAATCTCAAGCTTCAAGTGGATCGGATC 7901  
 Qy 1500 TCAGAGGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGAGTATCTGAATTTATGA 1559  
 Db 7902 TGAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGAGTATCTGAATTTATGA 7961

Qy 1560 TAAGTGTGTGATCATCTGACTGCAAAATATCCAACTCAGGCAATTCCTGACATGCGA 1619  
 Db 7962 TAAGTGTGTGATCATCTGACTGCAAAATATCCAACTCAGGCAATTCCTGACATGCGA 8021  
 Qy 1620 TACCAATATTTTTCAGGACCGAGTGAACCAAGCGGTTGTAACAGCCATATGACAGATC 1679  
 Db 8022 TACCAATATTTTTCAGGACCGAGTGAACCAAGCGGTTGTAACAGCCATATGACAGATC 8081  
 Qy 1680 GTCAAAATCTTCTCTGATGATGAGCAAAATTTCTGTTAGACATATCTCCGCTTTG 1739  
 Db 8082 GTCAAAATCTTCTCTGATGATGAGCAAAATTTCTGTTAGACATATCTCCGCTTTG 8141  
 Qy 1740 TGTCTCATGAGACGGGATGCAAAAGCAATGATTCCTGACATATGAGTGGAGATCC 1799  
 Db 8142 TGTCTCATGAGACGGGATGCAAAAGCAATGATTCCTGACATATGAGTGGAGATCC 8201  
 Qy 1800 GAATGTTAAAGTCTAGTGTCTCATGATTAATCATGTTGAGAGACGACTCCCAAGCTGTT 1859  
 Db 8202 GAATGTTAAAGTCTAGTGTCTCATGATTAATCATGTTGAGAGACGACTCCCAAGCTGTT 8261  
 Qy 1860 GGGGTCCTCCGACAAAGAGGATTAATGTCGTGCGGGAGCAATTTGAGATCTTT 1919  
 Db 8262 GGGGTCCTCCGACAAAGAGGATTAATGTCGTGCGGGAGCAATTTGAGATCTTT 8321  
 Qy 1920 CCGGCGGTTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCA 1979  
 Db 8322 CCGGCGGTTTGGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCA 8381  
 Qy 1980 ACATGATTTGGCTTACAGAGAGATCCGGGGAGCTTTCAAACTCAACCGGCTGTGA 2039  
 Db 8382 ACATGATTTGGCTTACAGAGAGATCCGGGGAGCTTTCAAACTCAACCGGCTGTGA 8441  
 Qy 2040 GATTTTATTTTGAAGAACTTTCTTCAAGCACTGACACGCGTAAATATACCGGAGT 2099  
 Db 8442 GATTTTATTTTGAAGAACTTTCTTCAAGCACTGACACGCGTAAATATACCGGAGT 8501  
 Qy 2100 TTACTTGGCGGGTGTGAGTGTCTTCAAGGTGATGGGTGAGGCGTATTCAGAC 2159  
 Db 8502 TTACTTGGCGGGTGTGAGTGTCTTCAAGGTGATGGGTGAGGCGTATTCAGAC 8561  
 Qy 2160 CCGGTGTAACCGCGTGTGTCATTAATCCAACTTGTGAGGCAATTTGGCAAGGCGCA 2219  
 Db 8562 CCGGTGTAACCGCGTGTGTCATTAATCCAACTTGTGAGGCAATTTGGCAAGGCGCA 8621  
 Qy 2220 TCCTCTGGAACACTCTTGAAGATATTAATCAACCGAGTAAGAAATTA 2267  
 Db 8622 TCCTCTGGAACACTCTTGAAGATATTAATCAACCGAGTAAGAAATTA 8669

RESULT 9  
 AB016260  
 LOCUS 206479 bp DNA circular BCT 12-JUN-2003  
 DEFINITION Agrobacterium tumefaciens plasmid pTI-SAKURA, complete sequence.  
 ACCESSION AB016260.1 AB007188 AB006857 AB006858  
 VERSION AB016260.1 GI:6498173  
 KEYWORDS  
 SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)  
 ORGANISM Agrobacterium tumefaciens  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

REFERENCE  
 1 Hattori, Y., Suzuki, K., Uraji, M., Ohta, N., Katoh, A. and Yoshida, K.  
 Genome structure of pTI-SAKURA (II): strategy for DNA sequencing of  
 a Japanese cherry-Ti plasmid  
 Nucleic Acids Symp. Ser. 37, 159-160 (1997)  
 MEDLINE 98247140  
 PUBMED 9586048  
 2 Hattori, Y., Suzuki, K., Ohta, N., Uraji, M., Katoh, A. and Yoshida, K.  
 Genome structure of pTI-SAKURA (V): Complete nucleotide sequence of  
 plasmid pTI-SAKURA's vir region in Agrobacterium tumefaciens  
 Nucleic Acids Symp. Ser. 39, 265-266 (1998)  
 JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3  
Ohta,N., Suzuki,K., Hattori,Y., Uraji,M., Katoh,A. and Yoshida,K.  
Genome structure of pTi-SAKURA (III): Characteristics of T-DNA  
Nucleic Acids Symp. Ser. 39, 185-186 (1998)

4  
Uraji,M., Suzuki,K., Ohta,N., Hattori,Y., Katoh,A. and Yoshida,K.  
Genome structure of pTi-SAKURA (IV): Characteristics of tra region  
Nucleic Acids Symp. Ser. 39, 187-188 (1998)

5  
Suzuki,K., Ohta,N., Hattori,Y., Uraji,M., Kato,A. and Yoshida,K.  
Novel structural difference between nopline- and octopine-type  
trf genes: construction of genetic and physical map and sequencing  
of trf/trai and rep gene clusters of a new Ti plasmid pTi-SAKURA  
Biochim. Biophys. Acta 1396 (1), 1-7 (1998)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

98193120  
9524202  
6  
Suzuki,K., Hattori,Y., Uraji,M., Ohta,N., Iwata,K., Murata,K.,  
Kato,A. and Yoshida,K.  
Complete nucleotide sequence of a plant tumor-inducing Ti plasmid  
Gene 242 (1-2), 331-336 (2000)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

20184752  
10721727  
7  
(bases 1 to 206479)  
Suzuki,K. and Yoshida,K.  
Direct Submission  
Submitted (22-JUL-1998) Kazuo Yoshida, Hiroshima University.

COMMENT On or before May 14, 2001 this sequence version replaced gi:2982659, gi:2982703.

```
FEATURES
source      Location/Qualifiers
1. .206479
```

```
gene join(205959,1,1435)
CDS /gene="tiorf195"
join(205959,1,1435)
```

|      |                          |   |
|------|--------------------------|---|
| gene | complement (1559. .2479) | . |
|      | /gene="tiorf1"           |   |
| CDS  | complement (1559. .2479) |   |

```

/translataion="MDVAPRVANWNNFNPIIVEEASAPAIKRTHTTFVKIVODVOCYALIA
TABEEMVPEYMLSGVWVPSGKPLPLDIAIGENPKRGDIIITNDPRATDVLTHMID
VTLLVYIPFDEBKALITCMATVAHSAIDIGAVPSSISAPATPEVOEGLRPEMKLYEGG
VLEAIKISKI'FOONDSI'IPTELMGDI'QAMI'SGKSNDRVSLCERVDYOSFEGMDVINA
LNEKRVANIRFTTIDGVEYSDVLEIGHGOLAPESVWTWVDSIEVD'FGTGPOLATA
AANYLVGATHTVPI'DICLLAAYLITVDP'PLTPNRSGLIRAHIRAGRVANVAPAGSG
GSRASATRAVDVITIGLLOALPEGLAAAGGQVAVSVASADPTGSGRVANVAVGIDT
GGGARGVCGVOSSEFSEYKORSEVPEVIEVUVMRLALVDPBPSRAAGFPASGAMLL
EIEBMBTNSNRAVTVTRVLRANFVPA'PMGFFGKEI'GLGLGKAVNNGPDRVSGKISVLE
LGQGDII'LRITS'TGGA'EEPLEDVAALAE'ENGLM'SEPARADVA'AVFPRDSKIDE
AATAEHHERENRANKSDDPNFCI'ROCODOWMS'QOTTRBELAS'PALTFEORIS'ISO
HHRLI'LAGKERYDVAKSDVPAVEERAGRL"
/complement(11559..2479)
/gene="t:torff1"
/complement(11559..2479)
/complement(11559..2479)
/gene="t:torff1"
/complement(11559..2479)
/gene="t:torff1"
/note="putative transcriptional regulatory gene"
/codon_start=1
/transl_table=1
/protein_id="BAA87626.1"
/db_xref="cdi:6498174"
/translation="MSQIITGIRCFVVAABEELHFRRAIRLINSQPTLSQOIKRLE
ELMVQVILSTRVAVLITRGASFPFHADIDILKIKITDEALNKLAEAGLSGGGELLIT
EAI'PSPATSOILPFLVDFRORPRTALIKEDVSI'IDEL'GIERGDFHGLRPPNSN

```

|      |  |
|------|--|
| gene | complement (2594. .3319)<br>/gene="tiorf2" |
| CDS  | complement (2594. .3319)                   |

```
/function="involved in nopaline transport"  
/note="Probable ncsm gene"
```

```

/transl_table=11
/protein_id="BAA87627.1"
/db_xref="GI:6498175"

```

gene  
complement (3321..4031)  
EMWLNPHNRARVGGTAAPKAETH"  
VWLIIKSTSLASTITIVETGATKAOLISATSPVEFIVAGAILTFTFVSRLLMLTA  
LALNPAATYSEIRIGSOVSIGQIEBARAVAGSTFLOFRRIIPPIAROLPAPYGENE  
KRIYISLALGYCVVITIRSTPLLVOMFELIYSSAGQPTFVSGVMSRSEPPCALILS  
/translation="MDIOLIIESPFLKLAAPPTLLTAFLISLIGPVSVPVVALMRIS

```

/ gene="tiorf3"
complement (3321. .4031)
CDS

```

```
/function="involved in nopaline transport"  
/note="Probable nocQ gene"
```

```

/transl_table=1
/protein_id="BAA87628.1"
/db_xref="GI:6498176"
/translation="MDLTLLQMDADGDELRGAMMTVVAAGYFPGIIIFGSLFPAAL
AKSRFMSLRIGQVAVTVVRVAPPELLIFLVPFGSLTRLRIANGLGSGYGEYEP
FIVCLITISAGVAVTVRAVLRVAPGQLEAKSKIGMSGLMLRLVRLIPDARFPA
LPGVGNWQVQLKDTLSTISVWGQVLEIMRFAAGASTQPFYFTTAAETLTLSSVS
LRGFLKAEKANGVDSQ"

```

|      |  |
|------|--|
| gene | complement (4106 . 4957)<br>/gene="tiorf4" |
| CDS  | complement (4106 . 4957)                   |

```

/function="involved in nopaline transport
(nopaline-binding protein)"

```

```
/note="Probable noct ge  
/codon start=1
```

```

/protein_id="BA876529.1"
/db_xref="GI:6498177"
/translation="MFPEFLNLAAAVVTGLAAAPTOAKYKSTITTEGSGVSYVYFNF
KDCAGSLIGDILDNDLCRNHIEAPEQWAGDIFSLTARADATLMAAGYOPAR
KEVAFSPRLPIIPMFLLTADSPLEKTDVAENLPLNITPEQGAELDKTPTFEGY
KFGVQATSTSEAHNKMMPQWPSSTDTYTDITDVAENLPLKGRIDASAPSPKPTDKE
DNKDLKMFGEGRMGTFEGKGVGIGIKEDDAKLFDAIDAAIADGTVQKLQSQWFG
"IDAAPKC"

```

|      |  |
|------|--|
| gene | complement (5022. .5795)                   |
| CDS  | /gene="trorf5"<br>complement (5022. .5795) |

```
/function="involved in nopaline transport"  
/note="Probable nosp gene"
```

```
/codon_start=1
/transl_table=11
```

```

/db_xref="GI:6498178"
/translation="MDATQPTLVADVHKNFGTLEVLKGISLTANKGDVSIIGSSGS

```

GKSTFLRCMNFLETPNKGRIAVG  
FNLWGHMTVLQNVMEGPLHLVKQ  
RVSIARALAMOPSALLFDEPTSA

gene  
VSSKVLFLKGGIETQGTPEVF  
6016. .6951

```

/ gene="tiorf6"
6016. .6951
CDS

```

```

/function="transcriptional regulator for nopaline
catabolism"
/note="Probable nocr gene"
/codon start=1

```

```
/transl_table=11
/protin_id="BAA87631.1"
/db_xref="GI:6498179"
/translation="MRPCVPMKAMTIOGROLEAPFAVMITGCMSTANVITTOPAI
SRIRLDEBIEIGISLPERGNRLPTREAGILEKYSRPHNGIOHIDRAAELEKXHM
GSLRVACYTAPALFMSGVYQTFIADRPDVSVDLTPVSTVELISLSDHLDGJSL
AGYSPGLTTEPVSFPAVCLPGRHLEDEYVHATDLBESLISPSVSLMOMDA
ALDSCVCHNRIESSLALNCDLVSRCMGVIGVDPFTADYSANPVIQRSFDPVIVY
HFAIVLPTDSEPPRLVSEFPRAALDLAKLPETI"
complement (7308. .8624)
/gene="clorf7"
/complement (7308. .8624)
/gene="clorf8"
/function="involved in conjugation"
/notes="Probable trb1 gene"
/codon_start=1
/transl_table=11
/protin_id="BAA87632.1"
/db_xref="GI:6498180"
/translation="MIOSLKLGTAQBEALSGMRIRLPIIIVVLVVAFLGVIFV
GLSRGLYPRGDEGPDTSNGSPASTADQIKRVDGIIIEPOLVTLLOTPVETOE
KRVNPFPTQGRREARPAQLESEVWPARIQROEBOYLREOROMARLQAND
AIDSPYAVDRDFRAGADTANASATNVCSTATTSGASDLYAALRAGLGQNDP
NGQSKDEFENADLKEGLVNPVPOQSTVELKRSVLPATLITGINSDLPGRTAQ
VSONVDSATGHRLLIPQGTLPGRYDSKVSFGOSRVLVMTDIIIPNSGLDQGM
GTDAGSGSRDVKRMLOTFPSAILIIVIGIDMAVPESSLTATODTASDAARN
FAETFRVADRTIORNDVQPTLEIRGVKFNVLVDQDIVFPGTYR"
complement (8637. .9113)
/gene="clorf8"
/complement (8637. .9113)
/gene="clorf8"
/function="involved in conjugation"
/notes="Probable trb1 gene"
```

Query Match 90.1%; Score 2044; DB 1; Length 206479;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 2141; Conservative 0; Mismatches 125; Indels 2; Gaps 2;

```
QY 1 ATGTGAGCTTCACTCTCTTGTATTAACGAGGAGATCATCTCCCAACCAAAATGATGAT 60
DB 188587 ATGTGAGCTTCACTCTCTTGTATTAACGAGGAGATCATCTCTCCCAAAATGATGAT 188646
QY 61 CTGACATGTCGATTAAGCGCGATGATGAGCCGCGAGGTTTCCGATGCTTTAGAA 120
DB 188647 CTGATATGTCGATTAAGCGCGATGATGAGCCGCGAGGTTTCCGATGCTTTAGAA 188706
QY 121 CGAAGACTTCTAGGCGAGAGGATTAATCAAGCTTCCACCGAGTGCAGCGCTGATTA 180
DB 188707 CGTGAAGCTTCTAGGCGAGAGGATTAATCAAGCTTCCGCGAGTGCAGCGCTGATTA 188766
QY 181 GCTTGAAGAGCGGCGGATGATGCTTCCCGAGATCTCAGTGTGGAAGGTAGCA 240
DB 188767 GCTTGAAGAGCGGCGGATGATGCTTCCCGAGATCTCAGTGTGGAAGGTAGCA 188826
QY 241 GTTCTCTCCGCTTATCTATTTGCAAGAAATTTGCGGCGGATGATGATCGAAA 300
DB 188827 GCTTCTCTCCGCTTATCTATTTGCAAGAAATTTGCGGCGGATGATGATCGAAA 188886
QY 301 CTTTGGCGCGGCGAGCAGTGAAGTGTCTGTTCCATTCGACTTGAACACCATTTGCATG 360
DB 188887 CTTTGGCGCGGCGAGCAGTGAAGTGTCTGTTCCATTCGACTTGAACACCATTTGCATG 188946
QY 361 GATTTTCCGAAGCAACTTAATCCAGCCCTGTTTTTGTGACGCGTAAAGATGCGCA 420
DB 188947 GATTTTCCGAAGCAACTTCTCCAAACCTGTTTTTGTGACGCGTAAAGATGCGCA 189006
QY 421 CCGATTGATCTTATGCTATTTGTCGCGATTTCAATCTCTAAGATGCGCGGTTTTCGAAAC 480
DB 189007 TCCAGCATCTTATGCTATTTGTCGCGATTTCAATCTCTAAGATGCGCGGTTTTCGAAAC 189066
QY 481 CTGCGCAATGCGCGCTGTACGAGATGCGACGATGAATGCGTTTACCGGAGTTTACCATTAAC 540
DB 189067 CTGCGCAATGCGCGCTGTACGAGATGCGACGATGAATGCGTTTACCGGAGTTTACCATTAAC 189126
```

```
QY 541 CTTGAAGGCGCGTGCATTTTGACATGATGATTAATGTCGAAACCTGATGCTGAAGGCT 600
DB 189127 CTTGAAGGCGCGTGCATTTTGACATGATGATTAATGTCGAAACCTGATGCTGAAGGCT 189186
QY 601 TCGGCAAGTTTCTTTCCAAACATTCAGCTTGTCTACGATCTGCAGACCGTTTTTGAACCA 660
DB 189187 TCGGCAAGTTTCTTTCCAAACATTCAGCTTGTCTACGATCTGCAGACCGTTTTTGAACCA 189246
QY 661 TGTTCGATAGTGAAGGATTCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 720
DB 189247 TGTTCGATAGTGAAGGATTCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGCG 189306
QY 721 GTCAATTGCGCTGCGCATTTCCGCACTCGTGTGCGCAAGCAAGCTGCTTCACTGCGAGTA 780
DB 189307 ATCATTTGGCGCTGCGCATTTCCGCACTCGTGTGCGCAAGCAAGCTGCTTCACTGCGAGTA 189366
QY 781 GACGATGTTACATATATATGAGCAAGTATCGTGTGAGGCAAGCTTTGTCAATGCT 840
DB 189367 GACGATGTTACATATATATGAGCAAGTATCGTGTGAGGCAAGCTTTGTCAATGCT 189426
QY 841 TTGAGGAGCGCTCTAGTGTGCGGCGCAAAATGGGGCGATGCGATTTCTCTGCTGCA 900
DB 189427 TTGAGGAGCGCTCTAGTGTGCGGCGCAAAATGGGGCGATGCGATTTCTCTGCTGCA 189486
QY 901 TTCTGCTTGTTTTCTTCTCGACGCTTACGCGCTGTCTGATGAGCGCGTTCCAAAT 960
DB 189487 TTCTGCTTGTTTTCTTCTCGACGCTTACGCGCTGTCTGATGAGCGCGTTCCAAAT 189546
QY 961 CCGGCAAGTGCACCTTACTTGTGTCTACCAAGCGCTCCAAATCATGTGGAAGCCGCG 1020
DB 189547 CCGGCAAGTGCACCTTACTTGTGTCTACCAAGCGCTCCAAATCATGTGGAAGCCGCG 189606
QY 1021 CAGCTGCACCGAGCGTTTCCATTCGCGTTTACAGCGTTGAGCGGTTCTTGAAGGAC 1080
DB 189607 CAGCTGCACCGAGCGTTTCCATTCGCGTTTACAGCGTTGAGCGGTTCTTGAAGGAC 189666
QY 1081 GCTTTTCATGAGCGAGATATGTTGTGCTTCCGCTGTGCTTATTAATCAAGCGCTTGA 1140
DB 189667 GCTTTTCATGAGCGAGATATGTTGTGCTTCCGCTGTGCTTATTAATCAAGCGCTTGA 189726
QY 1141 TCAGACACATTAAGTGGGCTCATGACTCTGCGCAAAATTTGGCTGAACGTTTCGGAGG 1200
DB 189727 TCAGACACATTAAGTGGGCTCATGACTCTGCGCAAAATTTGGCTGAACGTTTCGGAGG 189786
QY 1201 GAGTCTTCTCTCAGGATAGAGAGATCTTCTGCGCACACATCTCTGCTGCTGAT 1259
DB 189787 GAGTCTTCTCTCAGGATAGAGAGATCTTCTGCGCACACATCTCTGCTGCTGAT 189846
QY 1260 ACATGAGTCTTCTCATGATGATGGAAGCTTCAAGCTTAATGGAATGAGATCTGCGGG 1319
DB 189847 ACATGAGTCTTCTCATGATGATGGAAGCTTCAAGCTTAATGGAATGAGATCTGCGG 189905
QY 1320 GTTTGCTCAGTTTGAAGCGGCTTATTAAGATCTCTCGCTGCTGATCAACGATA 1379
DB 189906 GTTTGCTCAGTTTGAAGCGGCTTATTAAGATCTCTCGCTGCTGATCAACGATA 189965
QY 1380 TGAAGAAATCAACGAGATGCGCTGAAGGAATCTCAGAACTTCCAGTGGATGCGATC 1439
DB 189966 TGAAGAAATCAACGAGATGCGCTGAAGGAATCTCAGAACTTCCAGTGGATGCGATC 190025
QY 1440 TGAAGTGTAAACGATGTCTGTGAGCCAGCCATATGCGCATGTTCAATCGAGGCGAT 1499
DB 190026 TGAAGTGTAAACGATGTCTGTGAGCCAGCCATATGCGCATGTTCAATCGAGGCGAT 190085
QY 1500 TCAAGAGAAAGACAAATAAAGATTAAGGCTTAAAGCGGAGATCTGAATCTTTATGA 1559
DB 190086 TCAAGAGAGAAAGACAAATAAAGATTAAGGCTTAAAGCGGAGATCTGAATCTTTATGA 190145
QY 1560 TAAAGTGTGTCAATCTGGAATCTCGCAATATTCGAATGAGGATGCTGACATGCGA 1619
DB 190146 TAAAGTGTGTCAATCTGGAATCTCGCAATATTCGAATGAGGATGCTGACATGCGA 190205
```

|            |   |                    |  |                 |
|------------|---|--------------------|--|-----------------|
| QY         | 162   | TPCCAAATATTTTTT    | CAGGACCCAGGTGAACCAAGCGGTGTTAAAGACCAATATGACAGATC    | 1679            |
| Db         | 190206  | TACCCACATTTTTT     | CGTGACCAAGTGAACCAAGCGGTGTTAAAGACCAATATGACAGGCTC    | 190265          |
| QY         | 1680  | GTCAAAACCTCTT      | CCATGACTGTAAGAAATCTGGTTAGACCATATCTCCGCTTG          | 1739            |
| Db         | 190266  | GTCAAAACCTCTT      | CCATGACTGTAAGAAATTTGGTTAGACCATATCTCCGCTC           | 190325          |
| QY         | 1740  | TCTCTCATGAGACGGGAT | CGCAAAAGCAGTATATTCCTGGAACCTATGAGTCGAGATCC          | 1799            |
| Db         | 190326  | TCTCTCATGAGACGGGAT | CGCAAAAGCAGTATCTGCTGGAACCTATGAGTCGAGATCC           | 190385          |
| QY         | 1800  | GAATGTTAAAGGCTT    | AGTGTCTCATCGATTATACATGGAGAGACACTCCACAGCTGTT        | 1859            |
| Db         | 190386  | GAATGTTAAAGGCTT    | AGTGTCTCATCGATTATACATGGAGAGACACTCCACAGCTGTT        | 190445          |
| QY         | 1860  | GGCGGGTCCCCGACAA   | AAAAAGAGGATATATGTCGTGCGGGAGACGAATTTGAGATCTTT       | 1919            |
| Db         | 190446  | GGCGGGTCCCCGACAA   | AAAAAGAGGATATCTGTCTGCTGCGGAGACGAATTTGAGATCTTT      | 190505          |
| QY         | 1920  | CCGGGCGTTTGGCCAG   | ACCTATTTTCTGCGCTGCGCTGATTACGACCAAAATGTTATTCGA      | 1979            |
| Db         | 190506  | CCGGGCGTTTGGCCAG   | ACCTATGTTTCCGCGCTGCGCTGATTACGACCAAAATGTTTCA        | 190555          |
| QY         | 1980  | ACATGATTTGGCTTAC   | GACAGAGAAATCCGGGGGAGCTTTCAAACTGAACCCGCGTGTGA       | 2039            |
| Db         | 190566  | ACATGATTTGGCTTAC   | GACAGAGAAATCCGGGGGAGCTTTCAAACTGAACCCGCGTGTGA       | 190625          |
| QY         | 2040  | GGATTTTATATCTGAA   | GAACCTTTCTTTCAAGACCTGGAACAACCGCTAATGATACCCGAGT     | 2099            |
| Db         | 190626  | GGATTTTATATCTGAA   | GAACCTTTCTTTCAAGCGCTGGAACAATGATATGATACCCGAGT       | 190685          |
| QY         | 2100  | TTACTTGGCGGGTTG    | CGATTGTTCTTTCCTTCAACAGTGGATGGTGGAGGGTGTATTCAGAC    | 2159            |
| Db         | 190686  | TTACTTGGCGGGTTG    | CGATTGTTCTTTCCTTCAACCGTGGATGGTGGAGGGCGCTATTCAGAC   | 190745          |
| QY         | 2160  | CGCGTGTTAAACCG     | CGCTGTGTCAATTTCCACAAATTTGTGAGCGATTTTGGCAAAAGGGCAA  | 2219            |
| Db         | 190746  | CGCGTGTTAAACCG     | CGCTGTGTCAATTTCCACAAATTTGTGAGGATATTTTGGCAAAAGGACAA | 190805          |
| QY         | 2220  | TCCTCTCGAACACTCT   | GTGGAAGATATACACCGCACTAGAGAAATTA                    | 2267            |
| Db         | 190806  | TCCTCTCGAACACTCT   | GTGGAAGATATACACCGCAATAGAAATTA                      | 190853          |
| RESULT 10  |   |                    |  |                 |
| AF126446   |   |                    |  |                 |
| LOCUS      | AF126446  | 5968 bp            | DNA linear   | BCT 25-FEB-2000 |
| DEFINITION | Agrobacterium tumefaciens plasmid pTiC58 agroclonidine synthase (acd) gene, partial cds; and 5 protein (5), indole-3-acetamide hydrolase (iaah), and tryptophan monooxygenase (iaam) genes, complete cds. |                    |  |                 |
| ACCESSION  | AF126446  |                    |  |                 |
| VERSION    | AF126446.1  | GI:483565          |  |                 |
| KEYWORDS   | Agrobacterium tumefaciens (Rhizobium radiobacter)   |                    |  |                 |
| SOURCE     | Agrobacterium tumefaciens   |                    |  |                 |
| ORGANISM   | Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.   |                    |  |                 |
| REFERENCE  | 1   | (bases 1 to 5968)  |  |                 |
| AUTHORS    | Osten, L., Salomone, J. Y., Helfer, A., Schmidt, J., Hammann, P., and De Ruffray, P.  |                    |  |                 |
| TITLE      | Sequence and functional analysis of the left-hand part of the T-region from the nopaline-type Ti plasmid, pTiC58  |                    |  |                 |
| JOURNAL    | Plant Mol. Biol.  | 41 (6), 765-776    | (1999)   |                 |
| MEDLINE    | 20199402  |                    |  |                 |
| PUBMED     | 10737141  |                    |  |                 |
| REFERENCE  | 2   | (bases 1 to 5968)  |  |                 |
| AUTHORS    | Osten, L., De Ruffray, P., and Hammann, P.  |                    |  |                 |
| TITLE      | Direct Submission   |                    |  |                 |
| JOURNAL    | Submitted (08-FEB-1999) Cell Biology, IBMP, rue du General Zimmer 12, Strasbourg 67804, France  |                    |  |                 |
| FEATURES   | Location/Qualifiers   |                    |  |                 |

| source | 1..5968   |
|--------|---|
| gene   | /organism="Agrobacterium tumefaciens"   |
| CDS    | /mol_type="genomic DNA"   |
| gene   | /strain="C58"   |
| CDS    | /db_xref="taxon:358"  |
| gene   | /plasmid="pTc58"  |
| CDS    | /note="T1 plasmid, T-DNA"   |
| gene   | /complement(<1..19)   |
| CDS    | /gene="acs"   |
| gene   | /complement(<1..19)   |
| CDS    | /gene="acs"   |
| gene   | /codon_start=1  |
| CDS    | /transl_table=1   |
| gene   | /product="agrocinnopine synthase"   |
| CDS    | /protein_id="AAD30486.1"  |
| gene   | /db_xref="GI:4836566"   |
| CDS    | /translation="MFNLGK"   |
| gene   | 782..1462   |
| CDS    | /gene="5"   |
| gene   | 782..1462   |
| CDS    | /gene="5"   |
| gene   | /function="may catalyze synthesis of indole-3-lactate"  |
| CDS    | /codon_start=1  |
| gene   | /transl_table=1   |
| CDS    | /product="5 protein"  |
| gene   | /protein_id="AAD30487.1"  |
| CDS    | /db_xref="GI:4836567"   |
| gene   | /translation="MYSHRPIFIIDSNIDRRRLKVLRHTEIAYSPFOEDLIPAPQSSMNSITNTDVPIDPAIDEVKRFCEVACLLPGRAGLPIAIIINDSLITYYCSFQNAKRYAKHRFIDGSDGVAVISTVPYABGITEIKMSWHNNVCQNTSHTDLADYIAIAIPTSLQNPSPFSMKIGCDSFLAPSRVDFCWEIIAVGALPHDNLGKKRPVMSMAITFLRLRV"   |
| gene   | /complement(1953..3353)   |
| CDS    | /gene="1aah"  |
| gene   | /complement(1953..3353)   |
| CDS    | /gene="1aah"  |
| gene   | /function="synthesizes indole-3-acetic acid from indole-3-acetamide"  |
| CDS    | /codon_start=1  |
| gene   | /transl_table=1   |
| CDS    | /product="indole-3-acetamide hydrolase"   |
| gene   | /protein_id="AAD30488.1"  |
| CDS    | /db_xref="GI:4836568"   |
| gene   | /translation="WVPITSLAQTLERLRKRDYGLVKTILARCAKPLMLNLTATDMDGSRPAKTRDRHNGAGLGCGLPFCFKNIGTGPPTSAAPPALINHLPNIPSRPOTLSASAPLPASQNMHELRCGITSNNVATGAVRANPNPBLPGSGSGCAAAVASLMLGSGITDGTGASVRLPAPALCGVGFRTLARYPRDRIIPSPTRDITGIIAQCVADYIILDVIVSGRSKILPMPKGLRIGLPTTYVDLDADVAAPATTRLRLNRGTFFPFAEADIDHEILNLSGSLPIALVEFPRLAKKYLDDFVGFPLTFQRIISPDVANTVASAIDGHIISNDVEYLAROSPRPRLQATRYNYFRLQDLATLPTAPLAAKAIOESSVHINSGMNTPKRIYVRNVDPPSNAGLPGLSPACLTDPRLPVGMEIDGLAGSHRLAIGRALEKAINRSPDPAN"  |
| gene   | 3690..5957  |
| CDS    | /gene="1aah"  |
| gene   | 3690..5957  |
| CDS    | /gene="1aah"  |
| gene   | /function="synthesizes indole-3-acetamide from tryptophan"  |
| CDS    | /codon_start=1  |
| gene   | /transl_table=1   |
| CDS    | /product="tryptophan monoxygenase"  |
| gene   | /protein_id="AAD30489.1"  |
| CDS    | /db_xref="GI:4836569"   |
| gene   | /translation="WSASALINDNCDHSTKMVDLIWDKADDELDRVSDAFBSRSLASRGMTQISGBSAGLACKRLADGRFPETISAGEKYAALSAITYYGEKILGLISLSEPRPAAARVGLVAIDLAFPCMDFSBADLLQTLFLLSGRCASSDLSHFVAISISIKTARSRLLOMPYRSGKTRKVTGLTLEAVPFPDMVAYGNRLMKASGSPPTIDLYDYRSGKIDQSSGSGIGFPEDEVPRKVAIIAGAGISGLVNASLHAQVDVPTIYASDRGKIKASHAKRDSVVAEWGAMPAPASCLFPELERGLSGMRPFNGVDTNLVYOGKRLTMRKGGQOPKPLFHRVYSGMRAPLKGDFGHEGDIYASVATLTKSGDIPRASHDNDTYLNRFGSGSBSAIERIFLGTHPPGSGTNSFPHDMDLFLMGIGSGGFVFEVSGFTETLRVINGESBNOMRSEGISLPRPRLATVNVGVSQRIHVOYRALEKEXTKIKIRLTKSGHILIPSCVLMGDIKAAVYCDLEBQDPNGKGLVISTWEDSDSKLAVD |

KKERFCLJLDAISRSFPAPAOHLVPAACADYDQNVQHDWLTJDNAGAFKLNREBDF  
YSEHFEQALDNTNDVGVLYAGSCSFTGWVEGALQIATACNAVCALIIHNGGTLANDN  
PLEHSMKRYNRYRNRN"  
BASE COUNT 1543 a 1333 c 1472 g 1620 t  
ORIGIN

Query Match 90.1%; Score 2042.4; DB 1; Length 5968;  
Best Local Similarity 94.4%; Pred. No. 0;  
Matches 2140; Conservative 0; Mismatches 126; Indels 2; Gaps 2;

QY 1 ATGTCAAGCTTCACTCTCTCTTGTATACAGTGCATCATCTCCACCAAAATGTGGAT 60  
DB ATGTCAAGCTTCACTCTCTCTTGTATACAGTGCATCATCTCTACCAAAATGTGGAT 3749  
QY 61 CTGACATAGTGTCAATAGGCGGATGATGATGACCGGCGGTTTCCGATGCTTTTGA 120  
DB CTGATATATGTCATTAAGGCTGATGATGATGACCGGCGGTTTCCGATGCTTTTGA 3809  
QY 121 CGAAGAGCTTCTAGGGGAGAGGATTAATCAAGCTCCACCGAGTCAAGCGTGGTTA 180  
DB CGTGAAGCTTCTAGGGGAGAGGATTAATCAAGCTCCGCGAGTGCAGCGCTGGTTA 3869  
QY 181 GCTTGCAGAAAGCTGCGCATGCTGCTTCCCGAGATCTCACTGCTGAGAAAGTAGCA 240  
DB GCTTGCAGAAAGCTGCGCATGCTGCTTCCCGAGATCTCACTGCTGAGAAAGTAGCA 3929  
QY 241 GTTCTCTCCGCTTATATCTATATTGGCAAGAAATTTGGGGCGGATTAATTCGAAA 300  
DB GCCCTCTCCGCTTATATCTATATTGGCAAGAAATTTGGGGCGGATTAATTCGAAA 3989  
QY 301 CCTTGGCGCGGCGCAACAGTAGTGTCTGTTGCCATTCGACTTGAACACCATTTGCATG 360  
DB CTTTGGCGCGGAGAAAGTAGTGTCTGTTGCCATTCGACTTGAACACCATTTTGTATG 4049  
QY 361 GATTTCCTCGAAGCAACATTAATCAAGCCCTGTTTGTCTAGCGGTAAGATGTGA 420  
DB GATTTCCTCGAAGCAACATTAATCAAGCCCTGTTTGTCTAGCGGTAAGATGTGA 4109  
QY 421 CCGATTGATTTTGTATCTTTCGTGGCGATTTCAATCTCTAAGATGCGCGCTTTGCAAC 480  
DB TCACGACATCTTGTATCTTTCGTGGCGATTTCAATCTCTAAGATGCGCGCTTTGCAAC 4169  
QY 481 CTGCAATGCGCGTGTAGCAAGATGCGATGAATGCGTTACCGGTTTACCATTAAC 540  
DB CTGCAATGCGCGCTTTCGAGAAAGGACACGAAAGCGGTTTACCTTGAAC 4229  
QY 541 CTTGAAGGCGCGTGCATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 600  
DB CTTGAAGGCGCGTGCATTTGACATGATGATGATGATGATGATGATGATGATGATGAT 4289  
QY 601 TCGGCAAGCTTCTTTCGAAATTCGATGCTCTCAAGATGCGGTTTGTGACAA 660  
DB TCGGCAAGCTTCTTTCGAAATTCGATGCTCTCAAGATGCGGTTTGTGACAA 4349  
QY 661 TGTTCAGATGAGAGCGATGCGCTTCTTCGAGAGATGCTTCAAGCGGAAAGTGGC 720  
DB TGTTCAGATGAGAGCGATGCGCTTCTTCGAGAGATGCTTCAAGCGGAAAGTGGC 4409  
QY 721 GTCATTGGGCGCTGACATTTCCGAGCTGCTGCTGCAACGAACTGCTTCAATGCTGGGTA 780  
DB ATCATTTGGCGCTGACATTTCCGAGCTGCTGCTGCAACGAACTGCTTCAATGCTGGTA 4469  
QY 781 GAGAGTTTACATATATGAAAGAGATGCTGTTTGGAGGCAAGCTTTGTGACATGCT 840  
DB GAGAGTTTACATATATGAAAGAGATGCTGTTTGGAGGCAAGCTTTGTGACATGCT 4529  
QY 841 TTCAGGAGCGCTCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 900  
DB TTCAGGAGCGCTCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 4589  
QY 901 TTTGCTGTTGTTTCTTCTGAGCGCTTATCTTCTGATGAGCGGTTTCCCAAT 960  
DB TTTGCTGTTGTTTCTTCTGAGCGGCTTATCTTCTGATGAGCGGTTTCCCAAT 4649

QY 961 CCGGACAGTGCACCTTACTTGTCTACCAAGGCGCTCAATATATGATGAAAGCGGG 1020  
DB CCGGACAGTGCACCTTACTTGTCTACCAAGGCGCTCAATATATGATGAAAGCGGG 4709  
QY 1021 CAGCTGCACCGAAGCTGTTCCATCGCGTTTACCAAGGTTGGCGTCTTTTGAAGAC 1080  
DB CAGCTGCACCGAAGCTGTTCCATCGCGTTTACCAAGGTTGGCGTCTTTTGAAGAC 4769  
QY 1081 GGTTCATGAGGAGATATGTTGTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGAA 1140  
DB GGTTCATGAGGAGATATGTTGTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGAA 4829  
QY 1141 TCAGACACATTTAGGAGGCTCATGCTCTGCGCAATTTGGCTGAAACGTTTGGGAGG 1200  
DB TCAGAGACATTTAGGAGGCTCATGCTCTGCGCAATTTGGCTGAAACGTTTGGGAGG 4889  
QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTCTGTGTGT-AA 1259  
DB GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTCTGTGTGTAA 4949  
QY 1260 ACATGAGTTTCTCATGATTTGGGACCTATTTCAAGCTAATGGGAATAGATCTGGCGG 1319  
DB ACATGAGTTTCTCATGATTTGGGACCTATTTCAAGCTAATGGGAATAGATCTGGC-GG 5008  
QY 1320 GTTGGTCCAGTTTGAAGGCGGTTTATTTGAGATCTCTCGCTGCTGCTCAACGAGTA 1379  
DB GTTGGTCCAGTTTGAAGGCGGTTTATTTGAGATCTCTCGCTGCTGCTCAACGAGTA 5068  
QY 1380 TGAAGAAATTCAGCGGATGTCCCTGAGGAATCTCAGAACTTTCACGCTGATCGCATC 1439  
DB TGAAGAAATTCAGCGGATGTCTCTGAGGAATCTCAGAACTTTCACGCTGATCGCAC 5128  
QY 1440 TGAAGTGTAAAGGTTGTCTGTGAGCGGCAATATGCAATGTTCAATCGAGGCGAT 1499  
DB TGAAGTGTAAAGGTTGTCTGTGAGCGGCAATATGCAATGTTCAATCGAGGCGAT 5188  
QY 1500 TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGGCGGATTTCTGAACCTTATGA 1559  
DB TCAGAGGAAAGACAAAATTAAGATTAAGGCTTAAGAGGCGGATTTCTGAACCTTATGA 5248  
QY 1560 TAAAGTGTGTGCATCTGCACTGCAATATCAACTCAGGCAATTTGCTGACATGCGA 1619  
DB TAAAGTGTGTGCATCTGCACTGCAATATCAACTCAGGCAATTTGCTGACATGCGA 5308  
QY 1620 TACCAATTTTTTCAAGCACAGTGAACCAAGCGGTTGATTAAGGCTATGACAGATC 1679  
DB TACCAATTTTTTCAAGCACAGTGAACCAAGCGGTTGATTAAGGCTATGACAGATC 5368  
QY 1680 GTCAAAACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739  
DB GTCAAAACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5369  
QY 1740 TGTCTCATGAGCGGATGCGCAAAAGCATGATGCTGAGCTATGATGAGTGCAGATCC 1799  
DB TGTCTCATGAGCGGATGCGCAAAAGCATGATGCTGAGCTATGATGAGTGCAGATCC 5429  
QY 1800 GAATGTTAAAGGCTGATGCTCATGATTAATATGAGGAGAGAGATCTCCCAAGCTGTT 1859  
DB GAATGTTAAAGGCTGATGCTCATGATTAATATGAGGAGAGAGATCTCCCAAGCTGTT 5489  
QY 1860 GGGGATCCCGCAAAAAGAGGATTAATGCTGCGGAGCGCAATTTGAGATCTTT 1919  
DB GGGGATCCCGCAAAAAGAGGATTAATGCTGCGGAGCGCAATTTGAGATCTTT 5548  
QY 1920 CCGGCGTTTCCCAACACCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 1979  
DB CCGGCGTTTCCCAACACCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 5609  
QY 1980 ACATGATTTGCTTACAGACGAGATGCGGAGGAGCTTTCAACCTCAACCGGCTGTGTA 2039  
DB ACATGATTTGCTTACAGACGAGATGCGGAGGAGCTTTCAACCTCAACCGGCTGTGTA 5669

| Qy         | 2040  | GGATTTTATTCGTAAGAACTTTCTTCCAGCACTGCAACGCGTAAATGATACCGGAGT     | 2099                       |
|------------|---|---|----------------------------|
| Db         | 5729  | GGATTTTATTCGTAAGAACTTTCTTCCAGCGCTGACATGACATTAATGATACCGGAGT    | 5788                       |
| Qy         | 2100  | TTACTTGGCGGGTTGCAGTTGTTCTCTTACAGAGTGATGGGTGAGGGGTGCTAATTCACAG | 2153                       |
| Db         | 5789  | TTACTTGGCGGGTTGCAGTTGTTCTCTTACCGGTGATGGGTGAGGGCGCTAATTCACAC   | 5844                       |
| Qy         | 2160  | CGCGTGTAACGCCGCTCTGTGCAATTAATCCACAAATTGTGGAGGCATTTTGGCAAGGGCA | 2215                       |
| Db         | 5849  | CGCGGTGAACGCCGCTGTGTGCAATTAATCCAAATGTGGAGGATATTTTGGCAAGGACAA  | 5908                       |
| Qy         | 2220  | TCCTCTGCAACACTTTGGAAGAGATATTAATCAACCGCAGTGAATTA               | 2267                       |
| Db         | 5909  | TCCTCTGCAACACTTTGGAAGAGATATTAATCAACCGCAGTGAATTA               | 5956                       |
| RESULT 11  | AB025110  |   |                            |
| LOCUS      | AB025110  | 4660 bp   | DNA linear BCT 27-MAR-1999 |
| DEFINITION | Agrobacterium tumefaciens gene for indole acetamide hydrolase and tryptophan monooxygenase, complete cds.   |   |                            |
| ACCESSION  | AB025110  |   |                            |
| VERSION    | AB025110.1  | GI:4586311  |                            |
| KEYWORDS   | tryptophan monooxygenase; indole acetamide hydrolase.   |   |                            |
| SOURCE     | Agrobacterium tumefaciens (Rhizobium radiobacter)   |   |                            |
| ORGANISM   | Agrobacterium tumefaciens   |   |                            |
| REFERENCE  | Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Rhizobacterium; Agrobacterium group; Agrobacterium.   |   |                            |
| AUTHORS    | 1 (bases 1 to 4660)<br>Kasahara,T., Endo,S., Ebinuma,H., Sugita,K., Kawaoaka,A. and Morimae,N.  |   |                            |
| TITLE      | Agrobacterium tumefaciens plasmid PO22 indole acetamide hydrolase (iaah), tryptophan monooxygenase (iaam) genes   |   |                            |
| JOURNAL    | Published Only in Database (1999)   |   |                            |
| REFERENCE  | 2 (bases 1 to 4660)   |   |                            |
| AUTHORS    | Kasahara,T., Endo,S., Ebinuma,H., Sugita,K., Kawaoaka,A. and Morimae,N.   |   |                            |
| TITLE      | Direct Submission   |   |                            |
| JOURNAL    | Submitted (18-MAR-1999) Takehide Kasahara, Nippon Paper Industries Co., Ltd. R&D Div., Wood Bio-engineering Central Research Lab., 5-21-1, Oji, Kita-ku, Tokyo 114-0002, Japan  |   |                            |
| AUTHORS    | (E-mail:take@ntc.biglobe.ne.jp, Tel:81-3-3911-5106(ex.292), Fax:81-3-3914-3350)   |   |                            |
| FEATURES   | Location/Qualifiers   |   |                            |
| source     | 1..4660   |   |                            |
|            | /organism="Agrobacterium tumefaciens"   |   |                            |
|            | /mol_type="genomic DNA"   |   |                            |
|            | /strain="PO22"  |   |                            |
| gene       | /db_xref="taxon:358"  |   |                            |
|            | complement (353..1756)  |   |                            |
| CDS        | /gene="iaah"  |   |                            |
|            | complement (353..1756)  |   |                            |
|            | /gene="iaah"  |   |                            |
|            | /codon_start=1  |   |                            |
|            | /transl_table=1   |   |                            |
|            | /product="indole acetamide hydrolase"   |   |                            |
|            | /protein_id="BA076345.1"  |   |                            |
|            | /db_xref="GI:4586312"   |   |                            |
|            | /translation="MVPITLSAQLTERLRKRDYSGLEIVETLIRARCAAKLMLLATDWDLRRAKKNDRHGNAGLDCGILPCFKNIATGVPTSAATPALINHLPKIESRAERLFSAGALPGASGNHIELSGITSNNYVANGAVNPNPGLIPGSSGVAVASVASELMLGIGTGTGASVRLPALCGVVGFRPTLGRVPRDRIIPSPRTDVGIIAQCVAADVADIIILHOIVSGRSKASIPMLKGLRIGLPTTYVDADVAAPAEFTIRLRLANGTVAEADIPLEELNSGASIPALVEFPALAKKYLDDPVGVSPEDVKGIRSPVANI VSAOIDEHQJNSDEYELAROSFRRLQATNRYNRYLODAIILPTPLAKAIGQSSVYIHNSGNMFTFKIVKRVNDPSSNAGLPGSLPACTLPDRUPGVMEIDGLAGSHRLLAIGAALEKAINFTSFSPAPFN" |   |                            |
| gene       | 2096..4363  |   |                            |
|            | /gene="iaam"  |   |                            |
|            | 2096..4363  |   |                            |
| CDS        | /gene="iaam"  |   |                            |
|            | /codon_start=1  |   |                            |

[illegible]

|    |      |  |      |
|----|------|--|------|
| OY | 72   | GTCAATGGCCGTGCATTTCCCGACCTCGTGGTGGCAACGAACCTGCTCAATGCTGGGTA    | 780  |
| Db | 2816 | ATCATGTGGCGGTGGCATTTTCGGACCTCGTGGTACGACGAACGCTTCAATGCTGGTGA    | 2875 |
| OY | 781  | GACCATGTTACAAATATATGAAGCAAGTATCGTGTGGAGCAAGCTTTGGTCAATGCT      | 840  |
| Db | 2876 | GACCATGTTACAAATATATGAAGCAAGTATCGGGTGGAGCAAGCTTTGGTCAATGCT      | 2935 |
| OY | 841  | TTCAAGGACGCTCCTAGTGTGTGGCCGAAATGGGGGCGATGCGATTTCTCTGCTGCA      | 900  |
| Db | 2936 | TTCAAGAGTGTCTCCACGCGTGTGGCCGAATATGGGGGCGATGCGATTTCTCTGCTGCA    | 2995 |
| OY | 901  | TTCTGCTGTTTTTCTTCTCTGAGCGTTACGGCCCTGTCTTTGGATGAGGCCGTTCCCAAT   | 960  |
| Db | 2996 | TCGTGCTGTTTTTCTTCTCTGAGCGGTACGGCCCTGTCTTTGGATGAGGCCGTTCCCAAT   | 3055 |
| OY | 961  | CCCGGCAAGTCGACACTTACTTGGGTACCAAGGCGTCCAAATCATGTGAAAGCCGGG      | 1020 |
| Db | 3056 | CCCGGCAAGTCGACACTTACTTGGGTACCAAGGCTCCCAATCATGTGAAAGCCGGG       | 3115 |
| OY | 1021 | CAGCTGCCACCGAAGCTGTTCCATCGCGTTTACAACGGTTGGCGTTCGTTTGAAGAC      | 1080 |
| Db | 3116 | CAGCAGCCACCGAAGCTGTTCCATCGCGTTTACAGCGGTTGGCGTTCGTTTGAAGGAC     | 3175 |
| OY | 1081 | GCTTTTCAATGAGCGAGATATGTGTGGCTTCGCTGTGCTATTAATCTCAGGCTTGAA      | 1140 |
| Db | 3176 | GGTTTCCATGAGGGAGATATGTGTGGCTTCGCTGTGTTATTAATCTCAAGCCTTGAA      | 3235 |
| OY | 1141 | TCAGGACACATTAAGTGGGCTCATGACTCTCGGCAAAATTTGGCTGMAACGTTTGGGAG    | 1200 |
| Db | 3236 | TCAGGAGACATTAAGCGGGCTCATGACTCTCGGCAAACTTGGCTGMAACGTTTGGGAG     | 3295 |
| OY | 1201 | GAGTCTTCTCTTCAGGGATGAGAGATCTTTCGGGACACATCTCTGTGTGT-AA          | 1259 |
| Db | 3296 | GAGTCTTCTCTTCAGCGATGAGAGATCTTTCGGGACGACATCTCTGTGTGTGAA         | 3355 |
| OY | 1260 | ACATGGAATTTTCTCTCATGATTTGGGACCTATTCAAGCTAATGGAAATGGATCTGGCGG   | 1319 |
| Db | 3356 | ACATGGAATTTTCCCTCATGATTTGGGACCTAATCAAGCTAATGGAAATGGATCTGGCG-GG | 3414 |
| OY | 1320 | GTTTGGTCCAGTTTTTGAAGCGGGTTAATGAGATCCTCGGCTTGGTCATCAACGAGTA     | 1379 |
| Db | 3415 | GTTTGGTCCAGTTTTTGAAGCGGGTTAATGAGATCTTGGCTTGGTCATCAACGAGTA      | 3474 |
| OY | 1380 | TGAAGAAATCAGCGGATGTGCTCCCTGGAAGATCTCAGAACTTCCACGTCGGATCGCATC   | 1439 |
| Db | 3475 | TGAAGAAATCAGCGGATGTGCTCGTGAAGGATCTCAGAACTTCCACGTCGAATATGCTC    | 3534 |
| OY | 1440 | TGAAGTGTAAACGCTGTGTCTGTGAGCCAGCGCATATGCCATGTTCAAGTCAGGCGAT     | 1499 |
| Db | 3535 | TCAAGTGTAAACGCTGTGTCTGTGAACGACGATATAGCCATGTTCAAGTCAGGCGCAT     | 3594 |
| OY | 1500 | TCAGAAGAAAGCAAAATPAAAGTATAGGCTTAAGCGGGGATATGTGAACTTTATGA       | 1559 |
| Db | 3595 | TGAGAAGAAAGCAAAATPAAAGTATAGGCTTAAGCGGGGATATGTGAACTTTATGA       | 3654 |
| OY | 1560 | TAAAGTGTGTCACATCTGGACTCGCAAAATCAACTCAGGCAATGCTCGACATGCGA       | 1619 |
| Db | 3655 | TAAAGTGTGTCACTGTGACTCGCAAAATCAACTCAGGCAATGTCTGACATGCGA         | 3714 |
| OY | 1620 | TACCAATAATTTTTCAGGCAACGATGAACCAAGCGGTGATTAACAGCCATATGACGATC    | 1679 |
| Db | 3715 | TACCAACATTTTTCGTGCAACGATGAACCAAGCGGTGATTAACAGCCATATGACAGGCTC   | 3774 |
| OY | 1680 | GTCAAAATCTTTCCCTGATGACTGAACGAAATTCGTGTTAACCATATCCTCCGCTCTG     | 1739 |
| Db | 3775 | GTCAAAATCTTTTCGTGACTGAACGAAATTTGTGTTAACCATATCCTCCGCTCTG        | 3834 |
| OY | 1740 | TGTCCTCATGACCGGATCGCAAAAGCACTGATATGCTCGACATATGAGTGCAGAGATCC    | 1799 |
| Db | 3835 | TGTCCTCATGACCGGATCGCAAAAGCACTGTACTCTTGGACATATGAGCCGCAAGATCC    | 3894 |
| OY | 1800 | GAAATGTAAGGCTTAAGTCTCATCATGTTATATACATGGAGAGCATCCCAAGACTGTT     | 1859 |

|    |      |   |      |
|----|------|---|------|
| Db | 3695 | GAATGGTAAAGCTTGGTGGCCCCCACTTATPACATGGGAAGACGACTTCCACAAAGCTGTT   | 3954 |
| Oy | 1860 | GGCGGTCCTCCGACAAAAAAGNGCATATATGCTCTGTCGGGAGCGCAATTTTCGAGATCTTT  | 1919 |
| Db | 3955 | GGCGGTTCCCGACAAAAAAGAGGCAATTCTGTCTGTGGGGAGCGCAATTTTCGAGATCTTT   | 4014 |
| Oy | 1920 | CCGGCGCTTTGCCGACGACCTATTTCTGTGCTGCGCTGATTTAGACCAAAATGTTATTCA    | 1979 |
| Db | 4015 | CCGGCGCTTTGCCGAGATCTAGTTCCTGTGCTGCGCTGATTTAGACCAAAATGTTATTCA    | 4074 |
| Oy | 1980 | ACATGATTTGGCTTACAACGAGAAATGGCCGGGAGCTTTCAAATCTCAAACGGCGGTGGTGA  | 2039 |
| Db | 4075 | ACATGATTTGGCTTACAACGAGAAATGGCCGGGAGCTTTCAAATCTCAAACGGCGGTGGCGA  | 4134 |
| Oy | 2040 | GGATTTTATTTCTGAAGAACTTTTCTTCAAGCACTGGACACGGCTATGATACCGAGT       | 2099 |
| Db | 4135 | GGATTTTATTTCTGAAGAACTTTTCTTCAAGCGCTGGACATGCTATGATACCGAGT        | 4194 |
| Oy | 2100 | TTACTTTGGCGGGTTGCAGTGTCTCTTCAAGGTGATGGGTTGAGGGTGTATTTCAGAC      | 2159 |
| Db | 4195 | TTACTTTGGCGGGTTGCAGTGTCTCTTCAACGGTGGATGGGTTGAGGGTGTATTTCAGAC    | 4254 |
| Oy | 2160 | CGCGTGTAAAGCGCGTCTGTGCAATTATTCACAAATGTGGAGGCAATTTTGGCAAAAGGCGAA | 2219 |
| Db | 4255 | CGCGTGTAAAGCGCGTCTGTGCAATTATTCACAAATGTGGAGGTATTTTGGCAAAAGGCGAA  | 4314 |
| Oy | 2220 | TCCTCTGGAACACTCTTTGGAAGAGATTAATCTACCCGACAGTAGAAATT              | 2267 |
| Db | 4315 | TCCTCTGGAACACTCTTTGGAAGAGATTAATCTACCCGACAGTAGAAATT              | 4362 |

|            |   |
|------------|---|
| RESULT 12  |   |
| ATU237586  | ATU237586   |
| LOCUS      | 29524 bp DNA linear BCT 09-JUN-1999                             |
| DEFINITION | Agrobacterium tumefaciens Ti plasmid pTiC58 T-DNA region.       |
| ACCESSION  | AJ237586  |
| VERSIONS   | AJ237586.1 GI:5042179   |
| KEYWORDS   | acetoacetic synthase; agrocinopine synthase; IAH; isopenentenyl |

SOURCE  
transferrase; nopaline synthase; ornithine cyclodeaminase; protein 6a; protein 6b; torf1 gene; torf10 gene; torf11 gene; torf12 gene; torf13 gene; torf14 gene; torf15 gene; torf16 gene; torf17 gene; torf18 gene; torf19 gene; torf2 gene; torf20 gene; torf1 gene; torf12 gene; torf3 gene; torf4 gene; torf5 gene; torf6 gene; torf7 gene; torf8 gene; torf9 gene; tryptophan 2-monooxygenase. *Agrobacterium tumefaciens* (Rhizobium radiobacter)

|           |  |
|-----------|--|
| ORGANISM  | <i>Agrobacterium tumefaciens</i><br>Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;<br>Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium. |
| REFERENCE | 1  |

**AUTHORS** Gielen, J., Terry, N., Van Montagu, M. and Villarroel, R.  
**TITLE** Complete nucleotide sequence of the T-DNA region of the tumor inducing *Agrobacterium tumefaciens* Ti plasmid pTiC58 unpublished  
**JOURNAL** 2 (bases 1 to 29524)  
**REFERENCE**

**AUTHORS** Gielen, J.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (24-MAR-1999) Gielen J., Genetics, VIB/UG, KILedeganckstrat 35 Gent, 9000, BELGIUM  
**FEATURES** location/Qualifiers

```

SOURCE
1. .29524
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/db_xref="taxon:358"

```

```

gene      /plasmid="p1038"
          /complement(1..575)
          /gene="torf1"
CDS       /complement(<1..575)
          /gene="torf1"

```

```

/function="unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein"

```

```

gene
CDS
    /protein_id="CAB44625.1"
    /db_xref="GI:5042180"
    /translation="MPEDRRARGLRRAFLRLRFLCPSPHQAQVHRGDSPPRVANLF
    ROEIRRRPSRGQRTKIOEACLRMGQEBRBLSEBRAPFYAHADYICIFQALIIDGR
    OSINTFTFVDRLGKDHDPYAVRRKHGTAQARIAMRADAPSSRRWATAPMTISG
    RLVLVIGSATGSGMTNANSTGSSGAGRTS"
    complement(604.1479)
    /gene="torf2"
    complement(604.1479)
    /gene="torf2"
    complement(604.1479)
    /gene="torf2"
    complement(604.1479)
    /function="unknown"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44626.1"
    /db_xref="GI:5042181"
    /translation="WVQSPSGIGGRRAGATRPICEARFRTRNPTGRRRLQSGISASE
    LEIIMNNIATGFTASTDELDELRELRLEAGVTVELASLIDLJAGRIIVE
    RGLRLTLETFSPFRYVHGLVSSNFMDPATRRHQVDAKALVEICSIDARVLVOHA
    GFLRADQIARAGADEROCALLETGHAAYGIRIAPENIPTPEQVQOTPAEVAA
    TVKAVNHPDVALLDFKRLPGIQLSRPRPRSVARRHGVPILGAGRLSRRTATRLV
    VIIDGMECSGTWTS"
    complement(805.1353)
    /gene="torf3"
    complement(805.1353)
    /gene="torf3"
    complement(805.1353)
    /function="unknown"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44627.1"
    /db_xref="GI:5042182"
    /translation="MIDGFRGCGNFGSLTILARLGENILKGDANVTRGMAQPEQ
    RPAITVSAGALSDILCAQAKARVLDEAGVDVANLQCCISDILVPARCRVHRH
    EAVHIGAEGRLLKKGSEPIATLFDPAATCQIEAGKEFPRVDADCLKAQLIEAVK
    LAVRRCRIEARCNINVDPOF"
    complement(1359.1886)
    /gene="torf4"
    complement(1359.1886)
    /gene="torf4"
    complement(1359.1886)
    /function="unknown"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44628.1"
    /db_xref="GI:5042183"
    /translation="MPGEPAQOTALARLANVADKLWTGEIRTOYAAICYRTDSSGP
    IEVLITSRGTGVVPIKGMMAKKKHEVASQAEAGVGRVRRKKAHGYTVKR
    LGQGEFIPAMVOVHLIDVQRMEDDFPERHQRELQWSPRLAASAVGEBELRGLFARLE
    ERITQRAAVASKAG"
    complement(2090.2644)
    /gene="torf5"
    complement(2090.2644)
    /gene="torf5"
    complement(2090.2644)
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44629.1"
    /db_xref="GI:5042184"
    /translation="MGESGEIPGASGRVWVVIDPTFNVFVGQNMALISGLYENR
    RPSFGVIYAVRDILVGKTVPTPLNGOPIKPLPALMDSRASTGFSRFLSDLKISF
    RFGAATLSMIEVMGEBTDGVSIGDSTMVMALPLIGHLISNTIDMDRIDLSKTL
    RFGASBEFLSKRPLIEKVSAA"
    complement(3623.3647)
    /rpt_family="T-DNA left border"
    /rpt_type=DIRECT
    complement(4009.5124)
    /gene="torf6"
    complement(4009.5124)
    /gene="torf6"
    complement(4009.5124)
    /codon_start=1
    /transl_table=11
    /product="agrocinopine synthase"

gene
CDS
    /protein_id="CAB44630.1"
    /db_xref="GI:5042185"
    /translation="WYNLKGVECEHETSILTLALGEMERNLSELDITLISDLPVYS
    HDENIRISTLPLRLIRELSADVKNIPVIRREVSNGELTESITVNDITLPLEPLD
    KVDVPDGAATFFLDGRDFEAXHITLALSRPKRQVRLVLFYFEYSGGAFFDAVQ
    VSEBPMRETVLAMPVIFPOELPRLAKLIDLSIDIAVEDLYEGEKIWDMSLRQPMRY
    AVHIVMARVPEQLGCDKDEIVRAFNADYAAVRLAYVYVDNVEVRMRPHLKLATAT
    RCYDFSAOLENGEKAEPFINIRTRGPMREPRDERKXIRRGYGIIGNAAALADVVISDR
    SEDMSFWEMRNKVRRRDHCHPHLDVIEQDKSP"
    complement(6011.6694)
    /gene="torf7"
    complement(6011.6694)
    /gene="torf7"
    complement(6011.6694)
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44631.1"
    /db_xref="GI:5042186"
    /translation="MYERRPVFDIVDTWEIREREVLKPVLSVAEEDYCYFVONTVASA
    ORSWVLVANIPLNSPDSIDDALNRADAPCLHGPLLEPNLILKSPMYVYCSFEM
    RSCATRRPYREGISNRGVICTVPPYABGVTTRDMSWQONACVNTCSGRKDLAYIAF
    LPTSSLQESSYWHTKNGIYSFILAPSOTDAFCCEILCVGRALFEDRSRKEKLRNCLKL
    LNYRLKLP"
    complement(7150.8721)
    /gene="torf8"
    complement(7150.8721)
    /gene="torf8"
    complement(7150.8721)
    /function="unknown"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44632.1"
    /db_xref="GI:5042187"
    /translation="MENQTPVEALIERLKSSWLSLHVRKILPSPFLGPPYSYIE
    TVDQDADICIRGHSVTPLEDEGMVSHHQOIFCSHQPAAHDVDAFLRAABEAV
    IPDEDIADSVSPFASSENSLTDILMRMPLSELTANATQARKERDDAFVYLEPILP
    KEVKYLRSAOSIOTQEMRWOLKCHLPLMBRGLFWMLVIMQIAVLVDIYASAGA
    EVAVAPSGPITLOTOLMVAIIPYPADTRLAAHLACBVLGSLPVILIGLFPRIAPQ
    HEGTTRHTKADVAMKASLCLIMOMLQNRKIRDLKYSFLRKRIISTQSSDFTSEY
    TRLRNSITIIYSWLDTFEVSNRIRIQIGHKQDHEDLAVLRWGHIAODNDSPPGG
    KAGILVAIITMAGLICLSTFPDAYSGLAAANLVPLAARAADVMDNSQSHDLRL
    FTAAGVSPSTLFLVGNSSWMLTKKSLAGRANHFWIAFLIFPLSLPRLMGRL
    FMVYSGRLNDRSRA"
    complement(9762.10493)
    /gene="torf9"
    complement(9762.10493)
    /gene="torf9"
    complement(9762.10493)
    /function="phosphatase"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44633.1"
    /db_xref="GI:5042188"
    /translation="WHTSLMSYNONMIPSGDRPVPTIDLPANTSLEELQOGLCAAV
    ANYRDELTRKILPQOSWTVRVLSGDEETLRATKSCVFPCCSIRGAVODRIMO
    HDSPIITYSLNLSQYVEGKTLIRCYRFVLAASSLPRTIDIPABELKNVANSIIPG
    AYSMPRAVSCFVAIIPSKFTISMSKITTRGAEHEFVFSREENDIIDYIVAVKXT
    LFSQSDSEFGNLSPFDELONLIST"
    complement(11195.11890)
    /gene="torf10"
    complement(11195.11890)
    /gene="torf10"
    complement(11195.11890)
    /function="unknown"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="CAB44634.1"
    /db_xref="GI:5042189"
    /translation="MNVYQRPVTVYISLIRINREBAKVLIGAVVYVRNFWYEQTLAD
    AQKNWYKSLVLYDDPDGDAVTGILTFSRYACLAGRLGLDITAVANDPLIYICPRRK
    LEEYPERIVSFHSEIGSVCSMPDADGTVREKRYGHNLIISPGCLIPDLAEAVVA
    FLPSKSLKLPYSVYEVVHNDRYVHKFFALPSSRFHFVEVVAAGLAVPAKKKPPSGGI
    LACCTGKNTCL"

```

gene 12368..13312  
/gene="torf11"  
CDS 12368..13312  
/gene="torf11"  
/function="unknown"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="CAB44635.1"  
/db\_xref="GI:5042190"  
/translation="MNVPRPGRAPSLSEINDMEMHLFTSACILYHYIDNLLI  
YOEIWKYWRGLAESALKNELSLYLRLPPRPPTLQSLAEQEVLCKDAPLYVCS  
REMLSHVQSDKCOEGRANLACTLAPLYRTGLIEVRLCLPRLGNVETGDE  
DLSEFVALPPTAFRRGRPRVYLGENVQTFPGHILAFGLRSEILADSE  
SMWKSIVGESSRLNHVAVSELNRVPTAETDMMALGTDVAVVRPNLPPQD

Query Match 88.9%; Score 2016.2; DB 1; Length 29524;  
Best Local Similarity 94.2%; Pred. No. 0;  
Matches 2137; Conservative 0; Mismatches 128; Indels 4; Gaps 4;

QY 1 ATGTCAGCTTCACCTCTCTTGATTAACGAGCGGATATCTCCCAACCAAAATGGTGGAT 60  
Db 19987 ATGTCAGCTTCAGCTCTCTTGATTAACGAGCGGATATCTCTACCAAAATGGTGGAT 20046  
QY 61 CTGACAAATGATCGATTAAGCGGATGAATTGACCGCAGGGTTTCCAGTGCCTTTAGAA 120  
Db 20047 CTGATTAATGATCGATTAAGCGGATGAATTGACCGCAGGGTTTCCAGTGCCTTTAGAA 20106  
QY 121 CGAAGACCTTCTAGGGGAGAGAGATTACTCAAACTTCCACCGAGTCCAGCGCTGGGTTA 180  
Db 20107 CGTGAACCTTCTAGGGGAGAGAGATTACTCAAACTTCCCGCGAGTCCAGCGCTGGGTTA 20166  
QY 181 GCTTGAAGAGCGGCGGATGCTGCTCCCGAGATCTCAGCTGGTGAAGAGTGA 240  
Db 20167 GCTTGAAGAGCGGCGGATGCTGCTCCCGAGATCTCAGCTGGTGAAGAGTGA 20226  
QY 241 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTCGTGGGCGCGATCTGATCGAAA 300  
Db 20227 GCCTCTCCGCTTATATCTATTTGGCAAGAAATTCGTGGGCGCGATCTGATCGAAA 20286  
QY 301 CTTTGGGCGGCGGACAGAGTGTCTGTGTCATTCGACTTGGACACCATTTGCATG 360  
Db 20287 CTTTGGGCGGCGGACAGAGTGTCTGTGTCATTCGACTTGGACACCATTTGCATG 20346  
QY 361 GATTTCGCGAAGACAACTAATCCAGCCCTGTTTTGCGAGCGGTAAGATGCGA 420  
Db 20347 GATTTCGCGAAGACAACTTCTCAAACTTGTGTTGCGAGCGGTAAGATGCGA 20406  
QY 421 CCGATTGATCTAGTCAATTCGTGGCATTTCATCTTAAGATGCGCGCTTTCGAACC 480  
Db 20407 TCCAGCGATCTTAGTCAATTCGTGGCATTTCAGATCTTAAGATGCGCGCTTTCGAACC 20466  
QY 481 CTGCGAATGCGCGCTGTAAGAGATGCGACGATGAATGCGTTACCGGGTTTACATTAAC 540  
Db 20467 CTGCGAATGCGCGCTTTCGAGAAAGGCGACGAAACGCGTTACCGGGTTTACCTGACCC 20526  
QY 541 CTTGAAGGGGCGGCGCATTTGACATGAGTGGCTTAATGTTGAAACCTGATGCGAAGGT 600  
Db 20527 CTTGAAGAGCGCGCATTTGACATGAGTGGCTTAATGTTGAAACCTGATGCGAAGGT 20586  
QY 601 TCGGCAAGTTCCTTTCGAACATGCACTTGCTTACGATGCGACGCGTTTTTTTGAACAA 660  
Db 20587 TCGGCAAGTTCCTTTCGAACATTTGACTTGTCTATGACATAGGTTTTTTTGAACAA 20646  
QY 661 TGTTCGATAGTGAAGGATCGGCTTTCCTCGAGAGATGTTCTTAAGCGAAAGTGGCG 720  
Db 20647 TGTTCGATAGTGAAGGATCGGCTTTCCTCGAGAGATGTTCTTAAGCGAAAGTGGCG 20706  
QY 721 GTCAATGGCGCTGGCAATTCGCACTGTGTGGCAACGAACTGCTTCAATGCTGGGTA 780  
Db 20707 ATCAATGGCGCTGGCAATTCGCACTGTGTGGCAACGAACTGCTTCAATGCTGGGTA 20766  
QY 781 GACGATGTACAATATATGAAGCAAGTATGCTGTGGAGCAAGCTTTGTCAATGCT 840

Db 20767 GACGATGTACAATATATGAAGCAAGTATGCGGTTGAGGCAAGCTTTGTCAATGCT 20826  
QY 841 TTCAGGACGCTCTCTAGTGTGCGGCGGAAATGGGGGCGCATGGGATTTCTCTGCTGCA 900  
Db 20827 TTCAGGACGCTCTCTAGTGTGCGGCGGAAATGGGGGCGCATGGGATTTCTCTGCTGCA 20886  
QY 901 TTCGCTGTTGTTTTCTTCTCTGAGCGTTACGCGCTCTCTTCATGAGGCGGTTCCAAAT 960  
Db 20887 TCCTGCTGTTGTTTTCTTCTCTGAGCGGTTACGCGCTCTCTTCATGAGGCGGTTCCAAAT 20946  
QY 961 CCGGCAAGTGCACACTTACTTGTGTCACCAAGCGCTCCATATCATGTGAAAACCGGG 1020  
Db 20947 CCGGCAAGTGCACACTTACTTGTGTCACCAAGCGCTCCATATCATGTGAAAACCGGG 21006  
QY 1021 CAGCTGCCACCGGAGCTGTTCCATGCGGTTTACAAGGTTGGGCGGTTCTTGAAGGAC 1080  
Db 21007 CAGCAGCCACCGGAGGCTGTTCCATGCGGTTTACAGCGGTTGGCGGTTCTTGAAGGAC 21066  
QY 1081 GGTTCATGAGCGAGATATGTTGGGCTTCGCTGCTGCTATTTACTCAGGCTTGAAA 1140  
Db 21067 GGTTCATGAGGAGATATGTTGGGCTTCGCTGCTGCTATTTACTCAGGCTTGAAA 21126  
QY 1141 TCAGACACATTAGTGGGCTCATGACTCTGCGAAATTGGCTGAACGTTTCGGAGG 1200  
Db 21127 TCAGAGACATTAGGCGGCTCATGACTCTGCGCAACTTGGCTGAACGTTTCGGAGG 21186  
QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGCGCACATCTCTGCTGGT -AA 1259  
Db 21187 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTGCGCACATCTCTGCTGGTGGAA 21246  
QY 1260 ACATGAGATTTTCTCATGATTTGGGACTAATTCAGCTAATGGGAATAGGATCTGGCGG 1319  
Db 21247 ACATGAGATTTTCTCATGATTTGGGACTAATTCAGCTAATGGGAATAGGATCTGGCG -GG 21305  
QY 1320 GTTGTGTCAGTTTTGAAAACGGGTTTATGAGATCTTCCGTTGTCATCAACGAT 1379  
Db 21306 GTTGTGTCAGTTTTGAAAACGGGTTTATGAGATCTTCCGTTGTCATCAACGAT 21365  
QY 1380 TGAAGAAATTCAGCGGATGCTCCCTGAAGAAATCTGAGAACTTCAGCTGATCCATC 1439  
Db 21366 TGAAGAAATTCAGCGGATGCTCTGAAGAAATCTGAGAACTTCAGCTGATTCAC 21425  
QY 1440 TGAAGTGTAAACGAGTGTCTGTGAGCGACGATATGCAATGTTCAAGTCAAGCGAT 1499  
Db 21426 TCAAGTGTAAACGAGTGTCTGTGAGCGATATGCAAGTCAAGTCAAGCGAT 21485  
QY 1500 TCAGAGAAAGAAACAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATTAATGA 1559  
Db 21486 TGAAGAAAGAAAGCAAAATTAAGATTAAGGCTTAAGAGCGGATATCTGAATTAATGA 21545  
QY 1560 TAAAGTGTGTCACTGTGCACTCGCAAAATTCGCAACTCAGGCAATTCGCTGACATGCGA 1619  
Db 21546 TAAAGTGTGTGTCACTGTGCACTCGCAAAATTCGCAACTCAGGCAATTCGCTGACATGCGA 21605  
QY 1620 TACCAATATTTTTCAGAGCACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 1679  
Db 21606 TACCAACATTTTTCGAGCACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 21665  
QY 1680 GTCAAAATCTTCTCTGATGACTGAACGAAATTCGTTGATGACCATATCTCCGCTTTG 1739  
Db 21666 GTCAAAATCTTCTCTGATGACTGAACGAAATTTTGTGATGACCATATCTCCGCTCTG 21725  
QY 1740 TGTCTCTATGAGAGGATGCGCAAAACAGATGATTCCTCTGACTATGAGTCCGAGATCC 1799  
Db 21726 TGTCTCTATGAGAGGATGCGCAAAACAGATGATTCCTCTGACTATGAGTCCGAGATCC 21785  
QY 1800 GAATGTAAAGGCTAGTGTCTCATAGTTATATCATGAGAGAGCATCCCAACAGCTGTT 1859  
Db 21786 GAATGTAAAGGCTAGTGTCTCATAGTTATATCATGAGAGAGCATCCCAACAGCTGTT 21845  
QY 1860 GCGGATCCCGCAAAAGAGCATTAATGTCTGCTGGGAGCGCAATTTGAGATCTTT 1919

```
Db 21846 GGGGCTCCGACAAAAGAGCATTCGTCTGCTGCGGAGCAATTCGAGATCTTT 21905
Qy 1920 CCGGGCTTGGCCAGACCTATTTCTGCTGCGGATTAAGACCAAAATGTTATCA 1979
Db 21906 CCGGGCTTGGCCAGACCTATTTCTGCTGCGGATTAAGACCAAAATGTTATCA 21965
Qy 1980 ACATGATTTGGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACCGCGGTGTA 2039
Db 21966 ACATGATTTGGCTTACAGACGAGATGCGGGGAGCTTTCAAACTCAACCGCGGTGTA 22025
Qy 2040 GATTTTATTTATTCGAGAACTTTTCTTTCAGACATCGGACCTATATATACCGGAGT 2099
Db 22026 GATTTTATTTATTCGAGAACTTTTCTTTCAGAGCGCTGAGATCATATATATACCGGAGT 22085
Qy 2100 TTACTTGGCGGGTGTGAGTGTTCCTTCAAGGATGATGGTGAGAGGCTCTTTCAGAC 2159
Db 22086 TTACTTGGCGGGTGTGAGTGTTCCTTCAAGGATGATGGTGAGAGGCTCTTTCAGAC 22144
Qy 2160 CCGGTGTAACGCCGCTGTGCAATTTCCACAAATTTGTGAGGCAATTTGGCAAGGCA 2219
Db 22145 GCGGTGTAACGCCGCTGTGCAATTTCCACAAATTTGTGAGGCAATTTGGCAAGGCA 22204
Qy 2220 TCCTCT-CCGACACTCTTGGAGATATTAATTAATTAATTA 2267
Db 22205 TCCTCTACGAACTCTTGGAGATATTAATTAATTAATTA 22253

RESULT 13
ATTAAAH 10200 bp DNA linear BCT 06-JUL-2002
LOCUS A.tumefaciens Tm4 Ti plasmid DNA with TA-iaah interrupted by IS866,
DEFINITION TA-iaah, T-ipt and T-6 genes.
ACCESSION X56185.1 GI:39133
VERSION 1
KEYWORDS iaah gene; iaah gene; indoleacetamide hydrolase; ipt gene;
isopentenyl transferase; tryptophane monooxygenase.
SOURCE Agrobacterium tumefaciens (Rhizobium radiobacter)
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
REFERENCE 1 (bases 7402 to 8783)
AUTHORS Bonnard,G., Tindland,B., Paulus,F., Szegedi,E. and Otten,L.
TITLE Nucleotide sequence, evolutionary origin and biological role of a
rearranged cytokinin gene isolated from a wide host range biotype
III Agrobacterium strain
JOURNAL Mol. Gen. Genet. 216 (2-3), 428-438 (1989)
MEDLINE 89313683
PUBMED 2546041
REFERENCE 2
AUTHORS Bonnard,G., Vincent,F. and Otten,L.
TITLE Sequence of Agrobacterium tumefaciens biotype III auxin genes
JOURNAL Plant Mol. Biol. 16 (4), 733-738 (1991)
MEDLINE 91329707
PUBMED 1868204
REFERENCE 3 (bases 1 to 10200)
AUTHORS Bonnard,G.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1990) G. Bonnard, INSTITUT DE BIOLOGIE
MOLECULAIRE DES PLANTES DU CNRS, 12 RUE DU GENERAL ZIMMER, 67084
STRASBOURG CEDEX, FRANCE
FEATURES
source location/Qualifiers
1..10200
/organism="Agrobacterium tumefaciens"
/mol_type="genomic DNA"
/strain="Tm4"
/db_xref="taxon:358"
/plasmid="Ti"
/complement(join(192..653,3369..4319))
/gene="TA-iaah"
/complement(join(192..653,3369..4319))
/gene="TA-iaah"
/codon_start=1
/cranal_table=11

gene
CDS
BASE COUNT 2436 a 2471 c 2709 g 2584 t
ORIGIN
```

Query Match 82.9% Score 1880.8 DB 1: Length 10200;  
Best Local Similarity 89.9% Pred. No. 0;  
Matches 2039; Conservative 0; Mismatches 227; Indels 2; Gaps 2;  
QY 1 ATGTGACCTTACCTCTCTTGTATTAACAGTGGATCATCTCCCAACCAAAATGTGTGAT 60  
Db 4657 ATGTGACCTCATCTCTTGTATTAACAAATGGCATATTTCTCTACCAAAATGTGTGAT 4716  
QY 61 CTGACAAATGTGTGATTAAGGCGGATGAATTTGGACCGGAGGTTTCCATGCTTCTTGA 120  
Db 4717 CTGATATATGTGTGACCAAGGCTGATGAGTGGACCGCGCTTGGCGGACCTTCTCGAA 4776  
QY 121 CGAAGAGCTTCTAGGGGAGAGAGATTAATCAAGCTCCACCGAGTSCAGCGCTGGATTA 180  
Db 4777 CGAGAGCTTCTAGGGGAGAGAGATTAATCAAGCTCCACCGAGTSCAGCGCTGGATTA 4836  
QY 181 GCTTGCAGAAAGCTGGCGGATGTGCTGCTTCCGAGATCTCAGTGTGTGAGAAAGTAGCA 240  
Db 4837 GCTTGCAGAAAGCTGGCGGATGTGCTGCTTCCGAGATCTCAGTGTGTGAGAAAGTAGCA 4896  
QY 241 GTTCTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGCGGATTAATGATCGAAA 300  
Db 4897 GTCTCTCTCCGCTTATATCTATTTGGCAAGAAATTTGGGGCGGATTAATGATCGAAA 4956  
QY 301 CCTTGGGCGGGGCAACAGTGAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db 4957 GCTTGGGCGGGGCAACAGTGAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5016  
QY 361 GATTTCTCCGAGACAACTATTCAGACCCCTGTTTTTGTGAGCGGTAAAGATGTGCA 420  
Db 5017 GATATTTCCAGAGCTCAATCTTCTCAAAACATGAATTTGTGATGTGTAAGATGTGCA 5076  
QY 421 CCGATTGATCTTATGCTATTTCTGTGGCAATTTCAATCTCTAAGACTTCCGGCTTTTCAAC 480  
Db 5077 CCGAGCACTTATGCTATTTCTGTGGCAATTTCAATCTCTAAGACTTCCGGCTTTTCAAC 5136  
QY 481 CTGCGCAATGCGGCTGTGAGAAATGGCAAGTGAATGGCTTACCGGATTTACCATTAAC 540  
Db 5137 CTGCGCAATGCGGCTGTGAGAAATGGCAAGTGAATGGCTTACCGGATTTACCATTAAC 5196  
QY 541 CTGGAAGGGGCGGTGCTTGTGACATGTAGCTTATGTGCGAAACTGATGTCTGAAGGAT 600  
Db 5197 ATTGAAGGCAAGTACATTTGACATGTAGCTTATGTGCGAAACTGATGTCTGAAGGAT 5256  
QY 601 TCGGCAAGTCTTCTTCAACATGATCTGTCTGACATCTGACAGCCGTTTTTGAACCA 660  
Db 5257 TCGGCAAGTCTTCTTCAACATGATCTGTCTGACATCTGACAGCCGTTTTTGAACCA 5316  
QY 661 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGGATGTTCTTAAGCGGAAAGTGGCG 720  
Db 5317 TGTTCGATAGTGAAGGATCGGCTTCTTCCGAGGATGTTCTTAAGCGGAAAGTGGCG 5376  
QY 721 GTCATTTGCGCTGTGATTTCCGACCTGTGTGTGCAACGAACTGCTTCAATGTCTGGGTA 780  
Db 5377 GTCATTTGCGCTGTGATTTCCGACCTGTGTGTGCAACGAACTGCTTCAATGTCTGGGTA 5436  
QY 781 GAGCATGTACATATATGAAGCAAGTGTGTGTGAGGCAAGCTTTGTGTCAATGCT 840  
Db 5437 GAGCATGTACATATATGAAGCAAGTGTGTGTGAGGCAAGCTTTGTGTCAATGCT 5496  
QY 841 TTTCAGGAGCGCTCTAGTGTGTGTGTGCGGAAATGGGGGCGATGGATTTCCCTCGCTGCA 900  
Db 5497 TTTCAGGAGCGCTCTAGTGTGTGTGTGCGGAAATGGGGGCGATGGATTTCCCTCGCTGCA 5556  
QY 901 TTCTGCTGTTTTTTCTTCTCTGACGCTTACGCGCTGTCTTCTGATGAGCGGTTTCCCAAT 960  
Db 5557 TTCTGCTGTTTTTTCTTCTCTGACGCTTACGCGCTGTCTTCTGATGAGCGGTTTCCCAAT 5616  
QY 961 CCGGCAAGTGAAGCTTACTTGTGTGACCAAGCGCTCAATATCATGTGGAAGCCGGG 1020  
Db 5617 CCGGCAAGTGAAGCTTACTTGTGTGACCAAGCGCTCAATATCATGTGGAAGCCGGG 5676

QY 1021 CAGCTGCCACCGAAGCTTTCCATCGCTTTTACAAAGTGTGGCGCTTCTTGAAGAC 1080  
Db 5677 CAGACGCCACCGAAGTGTTTCCATCGCTTTACAGGGGGGCGATCGTTCTTGAAGAC 5736  
QY 1081 GGTTTTCATGAGCGAGATATTTGTGTGGCTTGGCTGTGCTTATTACTAGGCTTGA 1140  
Db 5737 GGTTTTCATGAGCGAGATATTTGTGTGGCGTGGCTGTGCTTATTACTAGGCTTGA 5796  
QY 1141 TCAGACACATTAAGGGGCTCATGACTCTGGCAAAATTTGGTGAACGTTTGGGAGG 1200  
Db 5797 TCAGGGACATTAAGGGGCTCATGACTCTGGCAAAATTTGGTGAACGTTTGGGAGG 5856  
QY 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACACATCTCTCTGTGG-TA 1259  
Db 5857 GAGTCTTCTCTTCAAGGATAGAGAGATCTTTCTGGGACGATCTCTCTGTGTAGAA 5916  
QY 1260 ACATGAGTTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAATAGATTTGGCGG 1319  
Db 5917 ACATGAGTTTTCTCATGATTTGGGACCTATTCAGCTAATGGGAATAGATTTGGC-GG 5975  
QY 1320 GTTTGGTCCAGTTTTTGAAGGGGTTTATGAGATCTCCGCTTGTGTCACAGGATTA 1379  
Db 5976 GTTTGGTCCAGTTTTTGAAGGGGTTTATGAGATCTCTCTGTGTGTCACAGGATTA 6035  
QY 1380 TGAAGAAATTCAGCGGATGTGCTTGAAGAAATCTCAGAACTTCAAGTGTGATC 1439  
Db 6036 TGAAGAAATTCAGCGGATGTGCTTGAAGAAATCTCAGAACTTCAAGTGTGATC 6095  
QY 1440 TGAAGTGTTAACGGTGTCTGTGAGCCAGGCAATATGCTCATGTTCAAGTCAAGGCGAT 1499  
Db 6096 TGAAGTGTTAACGGTGTCTGTGAGCCAGGCAATATGCTCATGTTCAAGTCAAGGCGAT 6155  
QY 1500 TCAGAGGAAAGGACAAATATTAAGATTAAGGCTTAAAGCGGGATATCTGAATTA 1559  
Db 6156 CGAGAGGAAAGGACAAATATTAAGATTAAGGCTTAAAGCGGGATATCTGAATTA 6215  
QY 1560 TAAAGTGTGTGTCATCTGCACTGCAATATCAACTCAGGATTTGCTTGAATGCA 1619  
Db 6216 TAAAGTGTGTGTCATCTGCACTGCAATATCAACTCAGGATTTGCTTGAATGCA 6275  
QY 1620 TACCAATATTTTCAAGGACAGTGAACCAAGCGTTGATTAAGCCATATGACAGATC 1679  
Db 6276 TACCAATATTTTCAAGGACAGTGAACCAAGCGTTGATTAAGCCATATGACAGATC 6335  
QY 1680 GTCAAACTCTCTGATGCTGAAGAAATTTCTGTTAGACATATCTCCGCTTGG 1739  
Db 6336 GTCAAACTCTCTGATGCTGAAGAAATTTCTGTTAGACATATCTCCGCTTGG 6395  
QY 1740 TGTCTCATGAGCGGATCGCAAAAGCAGTATTCCTGAGCTATGATGCGAGATCC 1799  
Db 6396 TGTCTCATGAGCGGATCGCAAAAGCAGTATTCCTGAGCTATGATGCGAGATCC 6455  
QY 1800 GAATGTAAAGTGTAGTGTCTCATGATTATCATGAGAGAGAGACTCCCAAGCTGT 1859  
Db 6456 GAATGTAAAGTGTAGTGTGTCTCATGATTATCATGAGAGAGAGACTCCCAAGCTGT 6515  
QY 1860 GGGGTCCTCCGACAAAGAGGATTAATGTCTGCTGGGAGCGCAATTTGAGATCTTT 1919  
Db 6516 GGGGTCCTCCGACAAAGAGGATTAATGTCTGCTGGGAGCGCAATTTGAGATCTTT 6575  
QY 1920 CCGGCGTTTTCCAGACACTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCA 1979  
Db 6576 CCGGCGTTTTCCAGACACTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATCA 6635  
QY 1980 ACATGATGTGCTTACAGAGAGATGCGGGGAGACTTTAACTCAACCGGCTGTGTA 2039  
Db 6636 ACATGATGTGCTTACAGAGAGATGCGGGGAGACTTTAACTCAACCGGCTGTGTA 6695  
QY 2040 GGAATTTTATTTGAGGAATTTTCTTCAAGCACTGGAACAGGCTTAATGATCCGGAGT 2099  
Db 6696 GGAATTTTATTTGAGGAATTTTCTTCAAGCACTGGAACAGGCTTAATGATCCGGAGT 6755  
QY 2100 TTAATTGGCGGGTGAAGTTGTTCTTTCACAGGTGATGGGTGAGGCTTATTCAGAC 2159

|||||  
Db 6756 TTACTTGCGGGGTGACGTTGCTTCTTACACCGGTGAGATGGGTGAGGCGCTATTTCAGAC 6815  
Qy 2160 CGCGTGAAGCCGCTCTGTAATTATCCAAATTTGAGGCAATTTGGCAAGGGCAA 2219  
Db 6816 CGCGTGAAGCCGCTCTGCGCAATTATCCAAATTTGGAGAGCGTTTGAAGGAGCA 6875  
Qy 2220 TCCTCTGACACCTCTGGAAGATATTAATACCGCAGTGAATTA 2267  
Db 6876 TCCTCTTAACACCCCTTGGAAGATATTAATACCGCAATTAATTA 6923  
RESULT 14  
AVU83987 15463 bp DNA linear BCT 02-MAR-2001  
LOCUS Agrobacterium vitis plasmid pT1m4 agrocinopine synthase and (5)  
DEFINITION genes, (1aah) pseudogene, and (1aam), (1pt), (6b) and (ocs) genes,  
complete cds.  
ACCESSION U83987  
VERSION U83987.1 GI:1814322  
KEYWORDS Agrobacterium vitis (Rhizobium vitis)  
ORGANISM Agrobacterium vitis  
REFERENCE 1 Agrobacterium vitis  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
1 (bases 1 to 15463)  
Osten, L. and De Ruffray, P.  
Sequence and functional analysis of the shoot-inducing octopine  
T-DNA of Agrobacterium vitis strain CG474  
Unpublished  
2 (bases 1 to 15463)  
Osten, L. and De Ruffray, P.  
Direct Submission  
Submitted (07-JUN-1997) Institute for Plant Molecular Biology, rue  
du General Zimmer 12, Strasbourg 67084, France  
LOCATION/Qualifiers  
1. 15463  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/strain="Tm4"  
/db\_xref="taxon:373"  
/plasmid="pT1m4"  
3857. 6588  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:373"  
/insertion\_seq="IS866"  
15438. 215463  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:373"  
/insertion\_seq="IS867"  
139. 15310  
/note="shoot-inducing octopine T-DNA"  
139. 165  
repeat\_region  
misc\_feature  
repeat\_region  
CDS  
complement (373. 1629)  
/codon\_start=1  
/transl\_table=11  
/product="agrocinopine synthase"  
/protein\_id="AAB41872.1"  
/db\_xref="GI:1814323"  
/translation="MEGTAPANTLHMDELPPVYLAHSHVPLKLERLENTLPLQ  
VIRHGFNIGNRIQECTASLLAALGQGRNISELDVCLTSDNIPVSHDLNTRVS  
EKLGKDFNEIHSSDINNVPVIRVSNGLIQRYRETIIDHILDEILGKVSANPD  
ATIFLDGRNEAIVAVMLSHRPVYQRAVILFTEFEPDGAFAVDVILTAQPSAMR  
KSIALMPALPEELCARLROVSEPTVDLVLAKAMVSLMODMRVAVAVPVSSE  
VSNRLDRAVVDKVLAFPSDQAAVLAHVKEDIMTKARPHLKRAAVTRCDPAAL  
LNGERAERESIDINTGRARPHETDERKHIMRKGTGNSKTIADWVTSRSEDEMAIW  
EMNQGIDREVSLSPLHDVVGTSSE".  
gene  
2235. 2918  
2235. 2918  
CDS

/gene="5"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAB41873.1"  
/db\_xref="GI:1814324"  
/translation="MDHSRPIFNVIDCSHIOQRRELEIVLLATRRAYLRFAPQDLPA  
QRSMIDSIIDINVIDPAIDVVRFRSVACLPAPYQMSLNTVLSMLVYCSFDM  
RRACRPFYEGSGKGVISTVPYQSGITKEMRCOMENGVCHITNNKNDLPAYTAF  
LPRTSIQNTSFNVKXIGGDAFLAPSRADPFCEIVAVGKALFQDNGLKAPKQMSMA  
LISLRLV"  
complement (3403. 7530)  
/gene="1aah"  
/gene="1aam"  
complement (3403. 3856, 6589. 7530)  
/note="interrupted by the bacterial insertion element  
IS866"  
/pseudo  
/codon\_start=1  
/transl\_table=11  
3857. 3864  
/note="flanking repeat, left border of the IS866 insertion  
element"  
/rpt\_type=direct  
6581. 6588  
/note="flanking repeat, right border of the IS866  
insertion element"  
/rpt\_type=direct  
7868. 10135  
/gene="1aam"  
7868. 10135  
/gene="1aam"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAB41874.1"  
/db\_xref="GI:1814325"  
/translation="MSASSLDKQCHPSTKIVLIMVDAKDELRVAAFGSERAS  
RERRISQISGECNAGLACRDLADRPDEISAGQVAVLVAITYVGEILLRWILEPES  
VTRVSGVATIDAPSCMDISRAQLQTNLUSGRKCARSDLSHPAITSERASRT  
LQAPYBEGSLKSVTGTIVIEAVPFDWVAGRNMLKASGSPFTIDLYVRLPL  
DKCSDSGRIGFPEEDVPKRVAVTGAIGLVAVSELHLAGVDVTVIYEGDVGGL  
WSHAFQDAPVVAEMGMRPPPAASCLFFLELYVGSNRPFPNPGTVDTLVYEGCR  
VMKAGQOPPKLFHRYVSGMHAFLKQGFLEGLIVLSPALTEALKSGDIRAHDSMO  
IWLNRGREGSFSATIERIFLTHPPGGETWSPPHMDLFLKMGISGGGPGVPSGFT  
EILRLVINYENQKRCSEGISLPRRIASOVYNGVSVCORIRHVOVRAIIEKRTIK  
IRKSGISELYDKVVTSGLANIQKHCLCTDTTIRAVNQAVIDSHHTGSKLPL  
TERKEWFDMLPSCVLMDGFAKAVYCLADVPQDPNGKGLVLSYNEDSHKLAAVAD  
KKEKRLCLRLDAISKSPFPAQHLVPACADYDQNVQHDMLTEBNAGRKLNKRGDF  
VSELEPFOALDITNDTGYVLAGSGSFTGQWEGAIQTACNVCATIIHCGGILANDN  
PLKHPMKRYNVRNRN"  
10824. 11546  
/gene="1pt"  
10824. 11546  
/gene="1pt"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAB41875.1"  
/db\_xref="GI:1814326"  
/translation="MDLRILFGPTCTGKTSTAIALAQDTGLPYLSIDRVQCCPOLSTG  
SGRPVYEBELGTTRLYLDLDRPLVKGIIIAEQAHERRLIABVYNEAAGLILEGSGSL  
LCKMQSGVWASDFRWHIIRHKLADEETFMKAARVKOMLCPAIGPSLIQELVLYMN  
EPRRLITKEIDGYRYAMLPASQNRITTPMLQLDADMGKLIHGAQVYLTHAROE  
HEPFPVSAAFGFGFPGPGAY"  
complement (12224. 12847)  
/gene="6b"  
complement (12224. 12847)  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AAB41876.1"  
/db\_xref="GI:1814327"  
/translation="MTVANMQRVDFTRILINAGELQRLQEARDFGALLAEIVYFHP  
GATPEGDDEYILITGGLVYVYVSEGTAROCALRLPLPNSNSNGFVTVAIPLMLNDT  
QTLNLTLORCDOGGI VNVYHGSRTNEPFLA LMLNSCFVRFGTDELINGASYGFIARG

gene  
complement (13779..14855)  
/gene="oc8"  
complement (13779..14855)  
/gene="oc8"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AB41877.1"  
/db\_xref="GI:1814328"  
/translation="MAKVATIGAGNLTALADLRIGTQPSINAPISNRSEFNDV  
CLASLEVEPDIGDFQPRLEDDLTGASGATFLTPVPLGQGLRELAKEKNSNS  
VVALPESATSLACKQTLVTPFAPLIVBSITSPVACRPAKAVLGVAFEEVAT  
QPLSEEVKGFLEVPNPQWOPHAPISFESNTNPVAPGLIARDISIEOGLIPK  
FYEPVPOAIRVIAIDEBRLSIDALGLESSTDFSKKWKYGHANREFEYEG  
VAIEPKTMNHRVLTEDVKHILVLMWEIAEALGVQVPEKSVVQASDVLNMLRT  
GRSLSLNIGGANAIVRALNGV"  
repeat\_region  
15287..15310  
/note="flanking repeat; right border of the T-DNA"  
BASE COUNT 3873 a 3665 c 3902 g 4023 t  
ORIGIN  
Query Match 82.9%; Score 1880.8; DB 1; Length 15463;  
Best Local Similarity 89.9%; Pred. No. 0;  
Matches 2039; Conservative 0; Mismatches 227; Indels 2; Gaps 2;  
Qy 1 ATGTCACTTACCTCTCTTGTATTAACGATGATCATCTCCCAACCAAAATGATGAT 60  
Db 7868 ATGTCACTTACCTCTCTTGTATTAACGATGATCATCTCTACCAAAATGATGAT 7927  
Qy 61 CTGACAAATGATGATTAAGGCGATGAATGACCGAGGGTTTCCGATGCTTTTGA 120  
Db 7928 CTGATATATGATGACAAAGGCTGATGATGATGACCGCGCGCTTCCGACGCTTCAGAA 7987  
Qy 121 CGAAGAGCTTCTAGGGGAGAGAGATTACTCAAAAGCTTCCAGAGTACAGCTGGTTA 180  
Db 7988 CGAAGAGCTTCTAGGGGAGAGAGATTACTCAAAAGCTTCCGAGAGTCAACGCTGGTTA 8047  
Qy 181 GCTTGCAGAAAGCTGCGCGATGCTGCTTCCCGAGATCTCAGCTGATGAGAAAGTGA 240  
Db 8048 GCTTGCAGAAAGCTGCGCGAGCTGCTTCCCGAGATCTCAGCGGTCAGAGAGGTGCA 8107  
Qy 241 GTTCTCTCGCTTATATCTATTTTGGCAAGAAATTTGGGGCGGATTAATTCGAAA 300  
Db 8108 GTTCTCTCGCTTATATCTATTTTGGCGAAGAAATTTCTGAGTGGATATCTGAAACGAA 8167  
Qy 301 CTTTGGGCGGGGCAACAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
Db 8168 GCTTGGGCGGGGCAACAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8227  
Qy 361 GATTTCTCCGAGAGCAACTAATCTCAAGCCCTGTTTGGCTGAGCGGTAAAGATGCA 420  
Db 8228 GATTTCTCCGAGAGCTCAACTTCTCCAAACATGAATTTGCTGATGATTAAGATGCA 8287  
Qy 421 CCGATGATCTTATGCTATTTCTGCGCCATTTCAATCTCTAAGACTGCGGCTTTGCAAC 480  
Db 8288 CCGAGGACCTTATGCTATTTCTGCGCCATTTCAATCTCTGAGACTGCGGCTTTGCAAC 8347  
Qy 481 CTGCGCAATGCGGCTGCTGAGAAATGCAAGATGAATGCGTTACCGGGTTTACATACC 540  
Db 8348 CTGCGCAATGCGGCTGCTGAGAAAGCTGCTGTAAGAAAGCTGTAAGCTGTAATC 8407  
Qy 541 CTGGAAGGCGCGTGGCACTTGTAGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 8408 ATTGAAGAGGCAATCACTTTGACATGATGATGATGATGATGATGATGATGATGAT 8467  
Qy 601 TCGGAGAGTTCTTTTCAACAATGCTGCTCTAAGCTGACAGACCGTTTGTGACCA 660  
Db 8468 TCGGAGAGTTCTTTTCAACAATGCTGCTCTAAGCTGACAGATTTGTTCTGACAA 8527  
Qy 661 TGTTCGATATGATGACGATGCTGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGG 720  
Db 8528 TGTTCGATATGATGACGATGCTGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGG 8587

Qy 721 GTCATATGCGCTGATGATTTCCGAGCTGCTGCTGCGCAACGAACTGCTTATGCGGAT 780  
Db 8588 GTCATATGCGGCTGATGATTTCCGAGCTGCTGCTGCGCAACGAACTGCTTATGCGGAT 8647  
Qy 781 GACGATGTTACATATATGAGCAAGTATGCTGTTGAGGCAAGCTTTGCTACATGCT 840  
Db 8648 GACGATGTTACATATATGAGCAAGTATGCTGTTGAGGCAAGCTTTGCTACATGCT 8707  
Qy 841 TTCAGGAGCGCTCTGATGCTGCGGCAATGAGGGGCGATGCGATTTCTCTGCTGCA 900  
Db 8708 TTCAGGAGCGCTCTGAGGCTGCTGCGGCAATGAGGGGCGATGCGATTTCTCTGCTGCA 8767  
Qy 901 TTTCTGCTTCTTCTTCTCTGAGCGTTAAGCGCTGCTTCTGATGAGCGCTTTCCAAAT 960  
Db 8768 TGTGCTTCTTCTTCTCTGAGCGATGAGCGCTGCTTCTGATGAGCGCTTTCCAAAT 8827  
Qy 961 CCGGCGACAGTGAACCTTATCTTGTGCTTCAAGCGGCTTCAATACATGTTGAAAGCCGG 1020  
Db 8828 CCGGCGACAGTGAACCTTATCTTGTGCTTCAAGCGGCTTCAATACATGTTGAAAGCCGG 8887  
Qy 1021 CAGCTGCAACGAGGCTTCTCATGCGGCTTCAAGCGGCTTCAAGCGGCTTCTTGAAGAC 1080  
Db 8888 CAGCTGCAACGAGGCTTCTCATGCGGCTTCAAGCGGCTTCAAGCGGCTTCTTGAAGAC 8947  
Qy 1081 GGTTCCTATGAGGAGATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 8948 GGTTCCTATGAGGAGATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9007  
Qy 1141 TCAGACACATTAAGTGGGCTCATGCTCTGCGCAATTTGGCTGAACCGTTTCCGGAG 1200  
Db 9008 TCAGGAGACATTAAGGCGGCTCATGCTCTGCGCAATTTGGCTGAACCGTTTCCGGAG 9067  
Qy 1201 GAGTCTTCTTCTTCAAGGATGAGAGATCTTCTGCGCAACATCTCTGCTGCTGCTGCT 1259  
Db 9068 GAGTCTTCTTCTTCAAGGATGAGAGATCTTCTGCGCAACATCTCTGCTGCTGCTGCT 9127  
Qy 1260 ACATGAGATTTTCTCATATGAGGATCTTCAAGCTTATGAGGATGAGATGAGATGAG 1319  
Db 9128 ACATGAGATTTTCTCATATGAGGATCTTCAAGCTTATGAGGATGAGATGAGATGAG 9186  
Qy 1320 GTTTCGCTGATTTTGAAGCGGCTTATGAGATCTCTGCTGCTGCTGCTGCTGCTGCT 1379  
Db 9187 GTTTCGCTGATTTTGAAGCGGCTTATGAGATCTCTGCTGCTGCTGCTGCTGCTGCT 9246  
Qy 1380 TGAAGAAATTCAGCGGATGCTGCTGAGAAATCTCAGAACTTCCAGTTCGATTCGAT 1439  
Db 9247 TGAAGAAATTCAGCGGATGCTGCTGAGAAATCTCAGAACTTCCAGTTCGATTCGAT 9306  
Qy 1440 TGAAGTGTAAAGCTGCTGCTGAGCAAGCGGATGATGCTGCTGCTGCTGCTGCTGCT 1499  
Db 9307 TGAAGTGTAAAGCTGCTGCTGAGCAAGCGGATGATGCTGCTGCTGCTGCTGCTGCT 9366  
Qy 1500 TCAGAGGAAAGCAAAATTAAGTAAAGCTTAAAGCGGATATCTGAATCTTTATGA 1559  
Db 9367 CGAGAGGAAAGCAAAATTAAGTAAAGCTTAAAGCGGATATCTGAATCTTTATGA 9426  
Qy 1560 TAAAGTGTGATGATGCTGCTGAGTTCGCAATATCAACTCAGGATTTGCTGATGCA 1619  
Db 9427 TAAAGTGTGATGATGCTGCTGAGTTCGCAATATCAACTCAGGATTTGCTGATGCA 9486  
Qy 1620 TACCAATATTTTCAAGGACCAAGTGAACCAAGCGGTTGATTAACAGCCATATGAAGATC 1679  
Db 9487 TACCAATATTTTCAAGGACCAAGTGAACCAAGCGGTTGATTAACAGCCATATGAAGATC 9546  
Qy 1680 GTCAAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739  
Db 9547 GTCAAAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9606  
Qy 1740 TGTTCCTATGAGAGGATGCAAGCAAGGATGATGATGATGATGATGATGATGATGAT 1799  
Db 9607 TGTTCCTATGAGAGGATGCAAGCAAGGATGATGATGATGATGATGATGATGATGAT 9666  
Qy 1800 GAATGTTAAAGGCTTATGCTCATCATGTTATATCATGAGAGAGCACTCCCAAGCTGTT 1859

Db 9667 GAAACGTAAGTCTGCTGCTCATAGTATATACATGGAGAGACACTCCACAAGCTATT 9726  
Qy 1860 GGGGTCCTCCGACAAAAGAGCATATATCTCTGCTCGGAGCCCAATTTTCAGATCTTT 1919  
Db 9727 GGGGTCCTCCGACAAAAGAGCATATATCTCTGCTCGGAGCCCAATTTTCAGATCTTT 9786  
Qy 1920 CCGGCGCTTGGCCAGCACCATTCTCTGCTCGGAGCTTACAGCAAAATGTTATTTCA 1979  
Db 9787 CCGGCGCTTGGCCAGCACCATTCTCTGCTCGGAGCTTACAGCAAAATGTTATTTCA 9846  
Qy 1980 ACATGATGCTTACAGACGAAATGCGGGGAGACTTTCAACTCAACCGCGTGTGTA 2039  
Db 9847 ACATGATGCTTACAGACGAAATGCGGGGAGACTTTCAACTCAACCGCGTGTGTA 9906  
Qy 2040 GGAATTTTATTTGAGAACTTTCTTTCAAGCACTGACACGCTATATGATACCGAGT 2099  
Db 9907 AGATTTTATTTGAGAACTTTCTTTCAAGCACTGACACGCTATATGATACCGAGT 9966  
Qy 2100 TTACTGCGCGGTTGACAGTGTCTCTTCAAGGATGATGGATGGAGGCTATTTACAG 2159  
Db 9967 TTACTGCGCGGTTGACAGTGTCTCTTCAAGGATGATGGATGGAGGCTATTTACAG 10026  
Qy 2160 CGCGTGTACGCGCTCTGTGCAATTAATCAACAATTTGTGAGGCAATTTGGCAAGGCAA 2219  
Db 10027 CGCGTGTACGCGCTCTGTGCAATTAATCAACAATTTGTGAGGCAATTTGGCAAGGCAA 10086  
Qy 2220 TCCTCTGAAACACTTTGAGAAAGATTAATCAATCCGACATGTAATTA 2267  
Db 10087 TCCTCTGAAACACTTTGAGAAAGATTAATCAATCCGACATGTAATTA 10134

## RESULT 15

LOCUS AVU83986 14960 bp DNA linear BCT 02-MAR-2001  
DEFINITION Agrobacterium vitis plasmid pTiCG474 agrocinopine synthase, (5),  
and (iaaH) genes, (iaaM) pseudogene, (ipt) and (6b) genes, and  
(lcs) pseudogene, complete cds.

VERSION U83986.1 GI:1814315

KEYWORDS Agrobacterium vitis (Rhizobium vitis)  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

REFERENCE 1 (bases 1 to 14960)  
Olsen, L. and De Ruffray, P.

TITLE Sequence and functional analysis of the shoot-inducing octopine  
T-DNA of Agrobacterium vitis strain CG474

UNPUBLISHED

2 (bases 1 to 14960)  
Olsen, L. and De Ruffray, P.

REFERENCE Direct Submission  
Submitted (07-JAN-1997) Institute for Plant Molecular Biology, rue  
du General Zimmer 12, Strasbourg 67084, France

FEATURES  
Location/Qualifiers

source

1..14960  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/strain="CG474"  
/db\_xref="taxon:373"  
/plasmid="pTiCG474"

source

7244..8395  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:373"  
/insertion\_seq="IS870-like insertion element"

source

11921..12767  
/organism="Agrobacterium vitis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:373"  
/insertion\_seq="IS869"

misc\_feature  
252..14745

repeat\_region

252..278  
/note="shoot-inducing octopine T-DNA"  
/note="flanking repeat; left border of the T-DNA"  
/rpt\_type=direct  
complement(496..1854)  
/codon\_start=1  
/transl\_table=11  
/product="agrocinopine synthase"  
/protein\_id="AA041866.1"  
/db\_xref="GI:1814316"

CDS

2363..3046  
/gene="5"  
2363..3046  
/gene="5"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AA041867.1"  
/db\_xref="GI:1814317"

gene

CDS

2363..3046  
/gene="5"  
2363..3046  
/gene="5"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AA041868.1"  
/db\_xref="GI:1814318"

gene

CDS

5306..8720  
/gene="iaaH"  
join(5306..7243,8396..8720)  
/gene="iaaM"  
/note="interrupted by a bacterial IS870-like insertion  
element"  
/pseudo  
/codon\_start=1  
/transl\_table=11  
7244..7247  
/note="flanking repeat; left border of the IS870-like  
insertion element"  
/rpt\_type=direct  
7301..8395  
/gene="ORF1"  
7301..8395  
/gene="ORF1"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AA041869.1"  
/db\_xref="GI:1814319"

gene

CDS

repeat\_region

gene

CDS

ABRLIRARQHWMTADMMATVAGISASAVRITWKAHGLAPHRWRQFKLSNDPKFVDX  
SADGVATVIMROTKSKTCWRQWERFAAECEGLRKTSPSRIVRGPDVAERIV  
EADIPHELDXXGASLPALVEFPALAKOYLDVPGTASFSDVINGRSDPDAIVKA  
OVDGHOISSEDEVELAROSFRPLOATYRNYFRLYOHDALFPTAPLAARAKICODFSVT  
HNGSTWNTFKIIVRVNVPSSNAGLPGLSLPVHLTPHRLPVGMEDIGLAGSDHRLAIG  
ALEKXIDHILFPNVLN"

LRDVGILYVDPAAHAIIVLSVDEKSOIQALDRTOPGLPMKKRGLGTWTHDYKRHGTTLT  
PALANVDGTVICGNMORRHOFIRFLKRIINQAPDDKSHVILINDYAAHKPKYRA  
MLERHPRFPHFTPTSCMLNAVEGFARLSKRRLKRGVHSHVVDLQAAINRLVBN  
KXKRPRTMADPDKIITIAAVRGHOLDISH"

repeat\_region  
/note="flanking repeat; right border of the 15870-like  
insertion element"  
/rpt\_type=direct  
9419..10141  
/gene="idp"  
9419..10141  
/gene="idp"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AAB41870.1"  
/db\_xref="ncbi:1814320"  
/translation="MDRLRFPGPTCTGKSTAIRLAQOTLPVLSLDRVOCROSTG  
SGRPVBEELKTRLYLEDRPLVKGIIIAKQAHRLIGEVYNTAEHGLILEGGSISL  
LRCAQSSVSTDFRMHIIIRKLADETFNNAKARVROMLRPAVPSIIQELVHLMN  
EPLRLILKEIDGVRVAMLPASQNIPTDMLQLDPDMGELIHGIAOEYLHAROE  
QEPFVSVAVEGEPFGMC"  
complement(10816..11439)  
/gene="6b"  
complement(10816..11439)  
/gene="6b"  
complement(10816..11439)  
/codon\_start=1  
/transl\_table=1  
/protein\_id="AAB41871.1"  
/db\_xref="ncbi:1814321"  
/translation="MTVANMQRVDLTIISAGEIISQRLLEGARTDFGALLDIDYFQPS  
VLSGERDEYVIFSGQELVYVYLRDIARCALRMLPSNSNFIATVAPMLMNA  
HGLNOYLDQRCDPGVLVNYGVPHNTOFLAIRNSNCFVPSADVINSESYGVARG  
NYTEBAGGDDVDEGEAGAAATTSQIQLNLYNPVALGSCQLSA"

gene  
complement(13214..14290)  
/gene="ocs"  
complement(13214..14290)  
/gene="ocs"  
/pseudo  
/codon\_start=1  
/transl\_table=1  
14719..14745  
/note="flanking repeat; right border of the T-DNA"

repeat\_region  
/rpt\_type=direct  
3510 c 3632 g 3958 t 6 others

BASE COUNT 3854 a 3510 c 3632 g 3958 t 6 others  
ORIGIN

Query Match 69.7%; Score 1581.4; DB 1; Length 14960;  
Best Local Similarity 88.8%; Pred. No. 0;  
Matches 1733; Conservative 0; Mismatches 216; Indels 2; Gaps 2;

QY 1 ATGTCAGCTCACCTCTCTTGATTAACGAGCGCATCTCCCAACCAAAATGATGAT 60  
DB 5306 ATGTCAGCTCACCTCTCTTGATTAACGAGCGCATCTCTTACCAAAATGATGAT 5365  
QY 61 CTGACATGTCGATTAAGCGGATGAATTGACCGCAGGGTTTCCAGTGCCTTTGAA 120  
DB 5366 CTGACATGTCGATTAAGCGGATGAATTGACCGCAGGGTTTCCAGTGCCTTTGAA 5425  
QY 121 CGAAGACTTCTAGGGAAGAGATTACTCAAACTCCACCGAGTCGACGGCTGTTA 180  
DB 5426 CGAAGACTTCTAGGGAAGAGATTACTCAAACTCCACCGAGTCGACGGCTGTTA 5485  
QY 181 GCTTGAAGAGCGGCGGATGTCGCTCCCGAGATCTCAGCTGGTGAAGAGTGA 240  
DB 5486 GCTTGAAGAGCGGCGGATGTCGCTCCCGAGATCTCAGCTGGTGAAGAGTGA 5545  
QY 241 GTTCTCTCCCTTATCTATTTGCAAAAGAAATTCGGGCGGATGATCGAAA 300  
DB 5546 GTTCTCTCCCTTATCTATTTGCAAAAGAAATTCGGGCGGATGATCGAAA 5605  
QY 301 CTTTGGGCGGGGCAACGATGATGTCGCTGCTCCATCGACTTGAACACCATTTGATG 360  
DB 5606 CCAATGGCGGAGCAAGATGATGTCGCTGCTCCATCGACTTGAACACCATTTGATG 5665

QY 361 GATTTCTCGAAGCACAACATAATCCAAAGCCCTGTTTCTGAGCCGTAAGATGCGA 420  
DB 5666 GATTTCTCGAAGCACAACATAATCCAAAGCCCTGTTTCTGAGCCGTAAGATGCGA 5725  
QY 421 CCGATGATCTTATGTCATTTCTGAGCCATTTCAATCTCAAGCTCCGCTTTCGAA 480  
DB 5726 CCGAGGACCTTATGTCATTTCTGAGCCATTTCAATCTCAAGCTCCGCTTTCGAA 5785  
QY 481 CTGCAATGCCGCTGACGAGATGACGATGAATGCGTTACCGGGTTTACATTAAC 540  
DB 5786 CTGCAATGCCGCTGACGAGATGACGATGAATGCGTTACCGGGTTTACATTAAC 5845  
QY 541 CTTGAAGGGGCGGTGACATTTTGAATGATGATGATGATGATGATGATGATGAT 600  
DB 5846 CTTGAAGGGGCGGTGACATTTTGAATGATGATGATGATGATGATGATGATGAT 5905  
QY 601 TCGGCAAGTTCCTTTCCAAACATGACTGCTCTACGACTGACAGCCGTTTTCGAA 660  
DB 5906 TCGGCAAGTTCCTTTCCAAACATGACTGCTCTACGACTGACAGATCGTTTTCGAT 5965  
QY 661 TGTTCGATAGTGAAGGATCGGCTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGC 720  
DB 5966 TGTTCGATAGTGAAGGATCGGCTCTTCCGAGAGATGTTCTTAAGCCGAAGTGGC 6025  
QY 721 GTCATTTGGGCGTGGCATTTCCGAGCTGATGATGATGATGATGATGATGATGAT 780  
DB 6026 GTCATTTGGGCGTGGCATTTCCGAGCTGATGATGATGATGATGATGATGATGAT 6085  
QY 781 GACGATGTTCAATATATGAGACGATGATGATGATGATGATGATGATGATGATGAT 840  
DB 6086 GACGATGTTCAATATATGAGACGATGATGATGATGATGATGATGATGATGATGAT 6145  
QY 841 TTCAGGAGCGCTCTGATGTCGAGCGCAAAATGGGGGCGATGCGATTTCTCTGCTGA 900  
DB 6146 TTCAGGAGCGCTCTGATGTCGAGCGCAAAATGGGGGCGATGCGATTTCTCTGCTGA 6205  
QY 901 TTCGCTTGTGTTTCTTCTCGAGCGTTACGCGCTGCTTCGATGAGCGGTTCCCAAT 960  
DB 6206 TTCGCTTGTGTTTCTTCTCGAGCGTTACGCGCTGCTTCGATGAGCGGTTCCCAAT 6265  
QY 961 CCGGCAAGTCGACACTTACTTGTGTCACAAAGCGTCCCAATCATGTGGAAGCCGG 1020  
DB 6266 CCGGCAAGTCGACACTTACTTGTGTCACAAAGCGTCCCAATCATGTGGAAGCCGG 6325  
QY 1021 CAGCTGCACCGAAGCGTTTCCATCGGTTTCAAGAGTTGGGCGTCTTGAAGAC 1080  
DB 6326 CAGCTGCACCGAAGCGTTTCCATCGGTTTCAAGAGTTGGGCGTCTTGAAGAC 6385  
QY 1081 GATTTTCATGAGCGAGATATGTTGTTGCTTCGCTGCTGCTTATTACTCAGGCTTGA 1140  
DB 6386 GATTTTCATGAGCGAGATATGTTGTTGCTTCGCTGCTGCTTATTACTCAGGCTTGA 6445  
QY 1141 TCAGACACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACGCTTTCGGA 1200  
DB 6446 TCAGACACATTAAGTGGGCTCATGACTCTGGCAAAATTTGGCTGAACGCTTTCGGA 6505  
QY 1201 GAGTCTTCTCTCAGGATGAGAGATCTTCTGGGCAACATCTCTCTGAGGT-AA 1259  
DB 6506 GAGTCTTCTCTCAGGATGAGAGATCTTCTGGGCAACATCTCTCTGAGGT-AA 6565  
QY 1260 ACATGAGTTTCTCATGATGATGAGACTTATCAAGCTAATAGGATGATGAGCGG 1319  
DB 6566 ACATGAGTTTCTCATGATGATGAGACTTATCAAGCTAATAGGATGATGAGCGG -GG 6624  
QY 1320 GTTGTGTCAGTTTGAAGCGGTTTATGAGATCTCCGCTGGTCTCAACGAGATA 1379  
DB 6625 GTTGTGTCAGTTTGAAGCGGTTTATGAGATCTCCGCTGGTCTCAACGAGATA 6684  
QY 1380 TGAAGAAAACACGAGATGTCCTGAAGGAATCTCAGAACTTCCAGTGGATGCGATC 1439  
DB 6685 TGAAGAAAACACGAGATGTCCTGAAGGAATCTCAGAACTTCCAGTGGATGCGATC 6744

QY 1440 TGAAGTGTAAAGTGTGTCTGTGAGCCAGCCATATGCCATGTTCAAGTCAGGGCAT 1499  
 Db 6745 TCAAGTGTAAAGTGTGTCTGTGAGCCAGCCATATGCCATGTTCAAGTCAGGGCAT 6804  
 QY 1500 TCAAGAGAAAAGACAAAATAAAGATAGGCTTAAAGCCGGATATCTGAACCTTATGA 1559  
 Db 6805 TGAGAAAGAAAAGACAAAATAAAGATAGGCTTAAAGCCGGATATCTGAACCTTATGA 6864  
 QY 1560 TAAGGTGTGTGATCAATCTGAGCTCCGAAATATCCAACTGAGCATTTGCTGACATGCCA 1619  
 Db 6865 TAAGGTGTGTGATCAATCTGAGCTCCGAAATATCCAACTGAGCATTTGCTGACATTTGA 6924  
 QY 1620 TACCAATATTTTTCAGGACCAAGTGAACCAAGGCTTGAATACAGCCATATGACAGGATC 1679  
 Db 6925 TACCACCATTTTTCGTACACCAAGTGAACCAAGGCTTGAATACAGCCATGACAGGCTC 6984  
 QY 1680 GTCAAAACTCTTCTGTGATGACTGAACGAAAATTCGTGTTAGACCATATTCCTCCGTTTG 1739  
 Db 6985 TTCAAAATGTTTGTGTGACTGAAACGAAAATTTTGGATCAACGATATGCTCCGTCCTG 7044  
 QY 1740 TGTCTCATGAGAGGATTCGCAAAAGAGTATGCTGAGCTATGAGTGCAGATCC 1799  
 Db 7045 TGTCTCATGAGAGGATTCGCAAAAGAGTATGCTGAGCTATGAGTGCAGATCC 7104  
 QY 1800 GAATGTTAAGTCTAGTGTCTCATCACTATATACATGAGAGAGCACTCCACAAAGCTGT 1859  
 Db 7105 GAACGGTAAAGTCTAGTGTCTCATCACTATATACATGAGAGAGCACTCTCAAAAGCTACT 7164  
 QY 1860 GGGGGTCCCGACAAAAGAGCGATTATGCTGCGGAGCGCAATTTGAGATCTTT 1919  
 Db 7165 GGGGGTCCCGACAAAAGAGCGATTATGCTGCTACGAGACGCAATTTGAAATCATTT 7224  
 QY 1920 CCGGGGTTTGCACAGACTATTTCTGCTGCC 1950  
 Db 7225 CCTGCGTTTGCCCGCATCTAGTGTACC 7255

Search completed: November 24, 2003, 01:37:08  
 Job time : 8374.46 secs

**THIS PAGE BLANK (USPTO)**



```

QY 181 GCTTGCAAAAGCTGCGCATGTCGCTTCCCGAGATCTCAGTGTGGAAAAGTAGCA 240
Db 181 GCTTGCAAAAGCTGCGCATGTCGCTTCCCGAGATCTCAGTGTGGAAAAGTAGCA 240
QY 241 GTTCTCTCCGCTTATATCTATATTGCGAAAATAATTCGTGGGCGGATTAATTGAATCGAAA 300
Db 241 GTTCTCTCCGCTTATATCTATATTGCGAAAATAATTCGTGGGCGGATTAATTGAATCGAAA 300
QY 301 CTTTGGGCGGGGCAAGAGAGTGTCTGTTGCCATCGACTTTGGACACCATTTGGATG 360
Db 301 CTTTGGGCGGGGCAAGAGAGTGTCTGTTGCCATCGACTTTGGACACCATTTGGATG 360
QY 361 GATTTCCTCCGAGACAATAATCAAGCCCTGTTTTGCTGAGCGGCTAAAGATGTCGA 420
Db 361 GATTTCCTCCGAGACAATAATCAAGCCCTGTTTTGCTGAGCGGCTAAAGATGTCGA 420
QY 421 CCGATGATCTTAGTCATTTCTGTGGCCATTTCATCTCTAAGACTGCGGCTTTGCAAC 480
Db 421 CCGATGATCTTAGTCATTTCTGTGGCCATTTCATCTCTAAGACTGCGGCTTTGCAAC 480
QY 481 CTGCGCAATGCCGCTGTACGAGATGCGACGATGAATGCGTTACCGGGTTTACATACC 540
Db 481 CTGCGCAATGCCGCTGTACGAGATGCGACGATGAATGCGTTACCGGGTTTACATACC 540
QY 541 CTTGAAGGGGCGGTCGCAATTTGACATGTAGCTTATGTGCGAAACCTGATGCTGAAGGT 600
Db 541 CTTGAAGGGGCGGTCGCAATTTGACATGTAGCTTATGTGCGAAACCTGATGCTGAAGGT 600
QY 601 TCGGCAAGGTCCTTTCCCAACATGCACTTGCTCTACGACTGCAAGCCGTTTTTGACCA 660
Db 601 TCGGCAAGGTCCTTTCCCAACATGCACTTGCTCTACGACTGCAAGCCGTTTTTGACCA 660
QY 661 TGTTCGCAATAGTGAACGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720
Db 661 TGTTCGCAATAGTGAACGATCGGCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCG 720
QY 721 GTCATTTGGCGCTGCAATTTCCGAGACTGCTGATGCGCAACGAACTGCTTCAAGTGGGTA 780
Db 721 GTCATTTGGCGCTGCAATTTCCGAGACTGCTGATGCGCAACGAACTGCTTCAAGTGGGTA 780
QY 781 GACGATGTTAATATATGAGCAAGATGATGTTGAGGCAAGCTTTGGTCAATGCT 840
Db 781 GACGATGTTAATATATGAGCAAGATGATGTTGAGGCAAGCTTTGGTCAATGCT 840
QY 841 TTCAGGAGCGCTCTAATGTCGTGCGGAAATGGGGCGGATGCGATTTCTCTGCTGCA 900
Db 841 TTCAGGAGCGCTCTAATGTCGTGCGGAAATGGGGCGGATGCGATTTCTCTGCTGCA 900
QY 901 TTTGCTGTTTTTTCTTCTCGAGCGTTACGCGCTGCTTCTGATGAGGCGGTTCCAAAT 960
Db 901 TTTGCTGTTTTTTCTTCTCGAGCGTTACGCGCTGCTTCTGATGAGGCGGTTCCAAAT 960
QY 961 CCGGCGCAGTGCACATTAATTGCTACCAAGCGCTTCAATATCATGTGAAAAGCCGGG 1020
Db 961 CCGGCGCAGTGCACATTAATTGCTACCAAGCGCTTCAATATCATGTGAAAAGCCGGG 1020
QY 1021 CAGTGCACCGAAAGCTGTTCCATCGGTTTACCAACGTTGGCGCTGCTTTGAAAGAC 1080
Db 1021 CAGTGCACCGAAAGCTGTTCCATCGGTTTACCAACGTTGGCGCTGCTTTGAAAGAC 1080
QY 1081 GGTTCATAGGAGATATGTTGTTGGCTTGGCTGCTGCTGCTATTAATCAAGCCCTTAAA 1140
Db 1081 GGTTCATAGGAGATATGTTGTTGGCTTGGCTGCTGCTGCTATTAATCAAGCCCTTAAA 1140
QY 1141 TCAGGACACATTAAGTGGGCTCATGATCTCGGCAAAATTTGGCTGAACCGTTTGGAGG 1200
Db 1141 TCAGGACACATTAAGTGGGCTCATGATCTCGGCAAAATTTGGCTGAACCGTTTGGAGG 1200
QY 1201 GAGTCTTCTCTTCAAGGATATGAGAGATCTTTCTGGGCAACATCTCTCTGTTGTTAAA 1260
Db 1201 GAGTCTTCTCTTCAAGGATATGAGAGATCTTTCTGGGCAACATCTCTCTGTTGTTAAA 1260
QY 1261 CATGAGTTTTTCTCATGATTTGGGACCTATTAAGCTAATGAGAAATGAGATCTGGCGGG 1320

```

```

Db 1261 CATGAGTTTTTCTCATGATTTGGGACCTATTAAGCTAATGAGAAATGAGATCTGGCGGG 1320
QY 1321 TTTGTCACATTTTGAAGCGGTTTATGAATCTCCGCTTGGTCAATCAACGATAT 1380
Db 1321 TTTGTCACATTTTGAAGCGGTTTATGAATCTCCGCTTGGTCAATCAACGATAT 1380
QY 1381 GAAGAAATCAGGAGATGTCCTGGAAGATCTCAAGACTTCAAGCTCGGATCGCATCT 1440
Db 1381 GAAGAAATCAGGAGATGTCCTGGAAGATCTCAAGACTTCAAGCTCGGATCGCATCT 1440
QY 1441 GAAGGTTAACGATGTCGTGTGAGCCAGCGATATGCCATGTTCAAGTCAGGCGATT 1500
Db 1441 GAAGGTTAACGATGTCGTGTGAGCCAGCGATATGCCATGTTCAAGTCAGGCGATT 1500
QY 1501 CAGAGAAAAGCAAAAATTAAGATTAAGCTTAAAGCGGATATCTGAACCTTTATGAT 1560
Db 1501 CAGAGAAAAGCAAAAATTAAGATTAAGCTTAAAGCGGATATCTGAACCTTTATGAT 1560
QY 1561 AAGGTGTGTCACTGTGCACTGCAATATCAACTGAGGATTTGCTGACATGCGAT 1620
Db 1561 AAGGTGTGTCACTGTGCACTGCAATATCAACTGAGGATTTGCTGACATGCGAT 1620
QY 1621 ACCAATATTTTTCAGGACCAAGTGAACCAAGCGTTGATTAACAGCATATGACAGATCG 1680
Db 1621 ACCAATATTTTTCAGGACCAAGTGAACCAAGCGTTGATTAACAGCATATGACAGATCG 1680
QY 1681 TCAAAATCTTCTGATGACTGAACGAAAATTTCTGTTAACAATCTCTCCGCTTGT 1740
Db 1681 TCAAAATCTTCTGATGACTGAACGAAAATTTCTGTTAACAATCTCTCCGCTTGT 1740
QY 1741 GTCCTCATGAGCGGAGATCGCAAAAGCAGTATTTGCTGTGACTATGAGTGCAGATCCG 1800
Db 1741 GTCCTCATGAGCGGAGATCGCAAAAGCAGTATTTGCTGTGACTATGAGTGCAGATCCG 1800
QY 1801 AATGTAAGTCTAGTCTCATCACTATTAATGAGAGACGACTTCCCAAGCTGTTG 1860
Db 1801 AATGTAAGTCTAGTCTCATCACTATTAATGAGAGACGACTTCCCAAGCTGTTG 1860
QY 1861 GCGGTCCCGCAAAAAGAGCATTAATGTCGTCTGCGGACCGCAATTTGAGATCTTTC 1920
Db 1861 GCGGTCCCGCAAAAAGAGCATTAATGTCGTCTGCGGACCGCAATTTGAGATCTTTC 1920
QY 1921 CCGGCGTTTCCAGACCTAATTTCTGCTGCGCTGATTAACGACCAAAATGTTATCAA 1980
Db 1921 CCGGCGTTTCCAGACCTAATTTCTGCTGCGCTGATTAACGACCAAAATGTTATCAA 1980
QY 1981 CATGATTTGCTTACAGACGAAATGCGGGGAGCTTTCAAATCAACCGGCTGTGAG 2040
Db 1981 CATGATTTGCTTACAGACGAAATGCGGGGAGCTTTCAAATCAACCGGCTGTGAG 2040
QY 2041 GATTTTATTTCTGAAGAACTTTCTTCAAGCACTGACACGCTATATGATCCGGATT 2100
Db 2041 GATTTTATTTCTGAAGAACTTTCTTCAAGCACTGACACGCTATATGATCCGGATT 2100
QY 2101 TACTTGGCGGTTGCAAGTTGTTCTTCAAGGAGATGGGTGAGGGGTGCTATTCAAGC 2160
Db 2101 TACTTGGCGGTTGCAAGTTGTTCTTCAAGGAGATGGGTGAGGGGTGCTATTCAAGC 2160
QY 2161 GCGGTGAACCGCTCTGTGCAATTAATCAAAATTTGAGAGCAATTTGGCAAGGGCAAT 2220
Db 2161 GCGGTGAACCGCTCTGTGCAATTAATCAAAATTTGAGAGCAATTTGGCAAGGGCAAT 2220
QY 2221 CTTCTGAAACCTTTGGAAGAGATTAATCAACCGCAGTGAATTTG 2268
Db 2221 CTTCTGAAACCTTTGGAAGAGATTAATCAACCGCAGTGAATTTG 2268

```

RESULT 2  
 US-10-187-339-9  
 ; Sequence 9, Application US/10187339  
 ; Publication No. US20030084478A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: Ream, Walt et al.
; TITLE OF INVENTION: Plants Having Enhanced Gall Resistance and
; TITLE OF INVENTION: Methods and Compositions for Producing the Same
; FILE REFERENCE: 53629
; CURRENT APPLICATION NUMBER: US/10/187,339
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US/09/434,837
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 60/107,185
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2544
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
; US-10-187-339-9

Query Match      78.9%; Score 1788.6; DB 14; Length 2544;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1791; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 12 ACCTCTCTTGATACGATGCGATCATCTCCCAACCAAAATGGTGATCTGACATGAT 71
DB 318 ACCTCTCTTGATACGATGCGATCATCTCCCAACCAAAATGGTGATCTGACATGAT 377
QY 72 CGATTAAGCGCGATTAATGGAACGCGAGGTTCCGATCCCTTCTAGAAAGAAAGCTTC 131
DB 378 CGATTAAGCGCGATTAATGGAACGCGAGGTTCCGATCCCTTCTAGAAAGAAAGCTTC 437
QY 132 TAGGGGAGAGGATTAATCTCAAAAGCTCCACGAGTGAGCGCTGGTTAGCTTGCAAAAG 191
DB 438 TAGGGGAGAGGATTAATCTCAAAAGCTCCACGAGTGAGCGCTGGTTAGCTTGCAAAAG 497
QY 192 GCTGCGCGATGTCGCTTCCCGAGATCTCAGCTGGTGAAGATGACAGTTCTCCCGC 251
DB 498 GCTGCGCGATGTCGCTTCCCGAGATCTCAGCTGGTGAAGATGACAGTTCTCCCGC 557
QY 252 TTATATCTATATTTGGCAAAAGAAATTTCTGGGGCGGATCTTGAACTTCCAACTTGGGGCG 311
DB 558 TTATATCTATATTTGGCAAAAGAAATTTCTGGGGCGGATCTTGAACTTCCAACTTGGGGCG 617
QY 312 GGGCAAGTGAAGTGTCTCGTTGCGATCGACTTGACACCATTTGACATGATTTCTCCGA 371
DB 618 GGGCAAGTGAAGTGTCTCGTTGCGATCGACTTGACACCATTTGACATGATTTCTCCGA 677
QY 372 AGCAACAATATCCAAAGCCCTGTTTTTGTGAGGGGTAAGATGTGCAACCGATTGATCT 431
DB 678 AGCAACAATATCCAAAGCCCTGTTTTTGTGAGGGGTAAGATGTGCAACCGATTGATCT 737
QY 432 TAATCATTTTGTGGCCATTTCAATCTCTAAGACTGCCGCTTTGAAACCTTCCCATGCC 491
DB 738 TAATCATTTTGTGGCCATTTCAATCTCTAAGACTGCCGCTTTGAAACCTTCCCATGCC 797
QY 492 GCTGTACGAGATGAGCAAGTGAATGCTTAACCGGGTTTACCATTAACCTTGAAGGGG 551
DB 798 GCTGTACGAGATGAGCAAGTGAATGCTTAACCGGGTTTACCATTAACCTTGAAGGGG 857
QY 552 CGTGCCATTTGACATGATGATTAATGTCGAAACCTGATCTGAAGGGTTCCGAGGTTTC 611
DB 858 CGTGCCATTTGACATGATGATTAATGTCGAAACCTGATCTGAAGGGTTCCGAGGTTTC 917
QY 612 CTTTCCCAACATGATGATGCTCTAAGACTGCAACCGTTTTTTTGAACCAATGTTCCGATAG 671
DB 918 CTTTCCCAACATGATGATGCTCTAAGACTGCAACCGTTTTTTTGAACCAATGTTCCGATAG 977
QY 672 TGAGACGATGCGGCTTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGCGGCTCAATTGGCG 731
DB 978 TGAGACGATGCGGCTTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGCGGCTCAATTGGCG 1037
QY 732 TGACATTTCCGAGCTCGTGTGTCGAAACGAATGCTTCACTGCTGGGTAGACGATGTTAC 791
DB 1038 TGACATTTCCGAGCTCGTGTGTCGAAACGAATGCTTCACTGCTGGGTAGACGATGTTAC 1097
```

```

QY 792 AATATATGAAGCAAGTATCGTGTGAGGCAAGCTTTGGTCAATGCTTTGAGGACGC 851
DB 1098 AATATATGAAGCAAGTATCGTGTGAGGCAAGCTTTGGTCAATGCTTTGAGGACGC 1157
QY 852 TCCTAGTGTCTGTCGCGCAAAATGGGGCGATGCGATTTCTCTCTGTCATTTCTGTTGT 911
DB 1158 TCCTAGTGTCTGTCGCGCAAAATGGGGCGATGCGATTTCTCTCTGTCATTTCTGTTGT 1217
QY 912 TTTCTTCTGAGCGTTACGCGCTGCTTGTGATGAGCGGTTCCAAATCCGCGACAGT 971
DB 1218 TTTCTTCTGAGCGTTACGCGCTGCTTGTGATGAGCGGTTCCAAATCCGCGACAGT 1277
QY 972 CGACATTTACTGTCGTACCAAGCGCTCCAAATACATGTGAAAGCCGGGACGTCGAC 1031
DB 1278 CGACATTTACTGTCGTACCAAGCGCTCCAAATACATGTGAAAGCCGGGACGTCGAC 1337
QY 1032 GAAGCTGTTCCATCGCGTTTACCAACGTTGGCGTGGCTTCTTGAAAGACGTTTTCATGA 1091
DB 1338 GAAGCTGTTCCATCGCGTTTACCAACGTTGGCGTGGCTTCTTGAAAGACGTTTTCATGA 1397
QY 1092 GCGAGATATTTGTTGGCTTCCGCTGTCCCTATTACTCAGGCTTGAATCAGACACAT 1151
DB 1398 GCGAGATATTTGTTGGCTTCCGCTGTCCCTATTACTCAGGCTTGAATCAGACACAT 1457
QY 1152 TAGTGGGCTCATGACTCCTGGCAAAATTTGGCTGAACCGTTTCCGGAGGAGTCCCTTC 1211
DB 1458 TAGTGGGCTCATGACTCCTGGCAAAATTTGGCTGAACCGTTTCCGGAGGAGTCCCTTC 1517
QY 1212 TTCAGGATAGAGAGATCTTCTGGGCAACATCTCTCTGTGTGTAACATGAGATTT 1271
DB 1518 TTCAGGATAGAGAGATCTTCTGGGCAACATCTCTCTGTGTGTAACATGAGATTT 1577
QY 1272 CCTCATGATTTGGAGCTTATTCAAGCTAATGGAATGATCTTGGCGGGTTGTCAGT 1331
DB 1578 CCTCATGATTTGGAGCTTATTCAAGCTAATGGAATGATCTTGGCGGGTTGTCAGT 1637
QY 1332 TTTTGAAGCGGGTTTATTTGAATCCTCGCTTGTGATCAACGGAATGAAGAAATCA 1391
DB 1638 TTTTGAAGCGGGTTTATTTGAATCCTCGCTTGTGATCAACGGAATGAAGAAATCA 1697
QY 1392 GCGAGTGTCCCTGAAGATCTCAGAACTTCCACGTCGATCGCATCTGAAGTGTAA 1451
DB 1698 GCGAGTGTCCCTGAAGATCTCAGAACTTCCACGTCGATCGCATCTGAAGTGTAA 1757
QY 1452 CGGTGTGTGTGAGCGAGCGATATGCGATGTTCAAGTCAGGGCGATTCAAGAGAAAA 1511
DB 1758 CGGTGTGTGTGAGCGAGCGATATGCGATGTTCAAGTCAGGGCGATTCAAGAGAAAA 1817
QY 1512 GACAAAAATTAAGATTAAGCTTAAAGCGGGATATCTGAACCTTTATGATTAAGTGTGT 1571
DB 1818 GACAAAAATTAAGATTAAGCTTAAAGCGGGATATCTGAACCTTTATGATTAAGTGTGT 1877
QY 1572 CACATCTGAGCTCGCAATATCAACTCAGGCAATTCCTGACATGCGATACCAATATTT 1631
DB 1878 CACATCTGAGCTCGCAATATCAACTCAGGCAATTCCTGACATGCGATACCAATATTT 1937
QY 1632 TCAGGACCAAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATGCTCAAACTCT 1691
DB 1938 TCAGGACCAAGTGAACCAAGCGGTGATTAACAGCCATATGACAGATGCTCAAACTCT 1997
QY 1692 CCTGATGATGAAACGAAATTTCTGTTTAAACCATATCTCTCCGCTTGTGTCTCATGGA 1751
DB 1998 CCTGATGATGAAACGAAATTTCTGTTTAAACCATATCTCTCCGCTTGTGTCTCATGGA 2057
QY 1752 CGGAGTCGCAAAAGCAGTATTTGCTGAGCATGAGTGCAGAGATCCGATGAT 1806
DB 2058 CGGAGTCGCAAAAGCAGTATTTGCTGAGCATGAGTGCAGAGATCCGATGAT 2112
```

RESULT 3  
US-10-187-339-7  
; Sequence 7, Application US/10187339



```

US-10-140-472-358
; Sequence 358, Application US/10140472
; Publication No. US2003013888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C168
; CURRENT APPLICATION NUMBER: US/10/140,472
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-472-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCGCGCATTTGACATGCTTGTGCGAAGCTGCGAAG 597
DB 298 HNSLSIQHVPRPFKNINKLQELDLSONFLAKEIGDKFLHFLPSLIQLDLSNFELOYVR 357
QY 598 GGTTCGCGAGGTCCTTCCACCAATCGACTGCTCAGACGACGCGTTTGTGAC 657
DB 358 ASMNLSQAASSLSKLILIRGVFEKLSFNLSPLHNLONLEVDLGNFIKIANLSWF 417
QY 658 CAATGTTCCGATAGTGAAGCGATCGGCTTTCCGAGAGATGCTTAAGCCGAAGTG 717
DB 418 KQFKRLKVIDLSVKNISPSGDSSEVGFCSNARTSVESYEPVLEQLHYFRYDKVARSCEF 477
QY 718 GCGGTGATTTGGCGCTGCGCATTTCCGACATCGTGTGGCAAGCAACTGCTTCACTGGG 777
DB 478 KKKEASFMVSNEBCYKRGQTLDSLNSKNSIFPVKSSDPQHSFLKCLNLSNLSQTLNGBE 537
QY 778 GTAGACGATTTGATCAATATATGAAGCAAGATCGTGTGAGGAGCGTTGGTCACAT 837
DB 538 FQPLAELRYLDPSNNRLDLHSTAFEBELHLEVLDISSNSHYFQSGITHMLNFTKNLKV 597
QY 838 GCTTTCAGGAGCGCTCTAGTGTGCGGCGAATGGGGCGATGCGATTTCCCTGCTGCT 897
DB 598 LQKLMNDDDISSTRTWESSELRLEFRGNHLDVLMBEGDNRVYLQFLKNLKLKEBLDI 657
QY 657 GCATTTGCTGTTTCTTCTTCCGACGCTTACGCGCTGCTTTCGATGAAGCGCGTTCCA 957
DB 658 SKNSISFLPSGVFDGPNLKNLSLAKNGLKFSFWMKLOCLKNLETLDSLHNOQLTVRPER 717
QY 958 AATCCGCGACAGTCACTGCTGCTACCAAGGCGTCAATCAATGGAAGCC 1017
DB 718 LSNCSLSLKNLILKNNQISLTYFLQDAPQLRYLDLSNFKIOMIKOTFFPENVLNLMK 777
QY 1018 GGGCAGCTCCCAAGAGCTTTCATCGCGTTTACCAAGCGTTGGCGCTTCTTGAAG 1077
DB 778 LLLHNRFLCTGDAVFVWVWVHTVETIPLATIDVTCVGBAHHKQGSVISLDTYTELDL 837

```

```

QY 1078 GACGGTTTCATGAGGAGATATTTGCTGCTGCGCTGCTTACTACTAGCGCTTG 1137
DB 838 TWLILFSLISVSILFLMVMWMTASHLYFMDVWYTHHCRAKINGYORLISPPDCYDAFIVY 897
QY 1138 AATCAGACACATATTAGTGAGGCTCATGACTCTGCGCAATTGCTGTAACCGTTTGGG 1197
DB 898 DTKDPAVTEWVLAELVAKLEDREKHFNLCLEBRLPQGPVLENLNLSQISLKKTVFWM 957
QY 1198 AGGAGCTCTTCTCTTCCGAGGATAGAGAGATCTTTTGGGACACAT 1245
DB 958 TDKVAKTENFKIAFYLSHQRLMDERKVDVILLIFLEKFPQSKFLQRLK 1005

RESULT 5
US-10-141-761-358
; Sequence 358, Application US/10141761
; Publication No. US20030148432A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C198
; CURRENT APPLICATION NUMBER: US/10/141,761
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 358
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-761-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCGCGCATTTGACATGAGCTTGTGCGAAGCTGCGAAG 597
DB 298 HNSLSIQHVPRPFKNINKLQELDLSONFLAKEIGAKFLHFLPSLIQLDLSNFELOYVR 357
QY 598 GGTTCGCGAGGTCCTTCCACCAATCGACTGCTCAGACGACGCGTTTGTGAC 657
DB 358 ASMNLSQAASSLSKLILIRGVFEKLSFNLSPLHNLONLEVDLGNFIKIANLSWF 417
QY 418 KQFKRLKVIDLSVKNISPSGDSSEVGFCSNARTSVESYEPVLEQLHYFRYDKVARSCEF 477
QY 478 KKKEASFMVSNEBCYKRGQTLDSLNSKNSIFPVKSSDPQHSFLKCLNLSNLSQTLNGBE 537
QY 537 GCATTTGCTGTTTCTTCTTCCGACGCTTACGCGCTGCTTTCGATGAAGCGCGTTCCA 957
DB 538 SKNSISFLPSGVFDGPNLKNLSLAKNGLKFSFWMKLOCLKNLETLDSLHNOQLTVRPER 717
QY 718 GCGGTGATTTGGCGCTGCGCATTTCCGACATCGTGTGGCAAGCAACTGCTTCACTGGG 777
DB 778 GTAGACGATTTGATCAATATATGAAGCAAGATCGTGTGAGGAGCGTTGGTCACAT 837
DB 837 FQPLAELRYLDPSNNRLDLHSTAFEBELHLEVLDISSNSHYFQSGITHMLNFTKNLKV 597

```

QY 838 GCTTTCAGGAGCGCTCTAGTGTGTCGGCGAAATGGGCGCATGGCATTTCTCTGCT 897  
DB 598 LQKLMNNDISSISRTWSESLRTLBERGNHLDVMBGDNRYQLFRNLKLELDI 657  
QY 898 GCATTCTGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTTCGATGAGCGCGTTCCA 957  
DB 658 SKNSLSFLPSGVFDGMPNPKNLKSLAKNGLKSPWKKLOCLKNLETLDSHNLTVPER 717  
QY 958 AATCCCGGACAGTCACACTTACTTGCTACCAAGCGCTCCAAATACATGTGGAAAGCC 1017  
DB 718 LNSCSRLKNLILKNNQIRSLTYFLQDAFOLRYLDLSNKKIQIKTSPENVNLNLM 777  
QY 1018 GGGCAGCTGCGACCGAGCGTTCATCGGCTTACACGCGTGGCGCTTCTGAAG 1077  
DB 778 LLIHNRFLCTCDVAFVWVWVNTTEVTIPLATDVTCVPGAHKGSVISLDTYBELD 837  
QY 1078 GACGCTTTTCATGAGAGATATTGTGTGGCTTCGCTGCTGCTATTAATCAGCGCTTG 1137  
DB 838 TNLILFSLISVSLFLMWMNTASHLYFWDVWYIYHFCXAKIKGYORLISPDCCYDAFIY 897  
QY 1138 AATCAGACACATTAGTGCGCTCATGACTCTCGGCAATTGGCTGAACGCTTCCGG 1197  
DB 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLBERDWLPQPVLENTSOSIQLSKTVFWM 957  
QY 1198 AGGAGTCTCTCTTCAGGATAGAGAGATCTTCTGGGACACAT 1245  
DB 958 TDKYAKTENFKIAFYLSHQLMDEKVDVILLIFLEKPFQSKFLQLRK 1005

## RESULT 6

US-10-142-885-358  
; Sequence 358, Application US/10142885  
; Publication No. US20030157604A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Geriltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C248  
; CURRENT APPLICATION NUMBER: US/10/142,885  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-142-885-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;  
Best Local Similarity 5.4%; Pred. No. 0.028;  
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAGGGCGCGCGCATTTGACATGCTAGCTTATGCTGAACCTGATCTGAAG 597  
DB 298 HSNLSLQHPVPRFWFNKINLQELDSQNLAKELIDDAKFLHPLPSLIQDLSFNELQYRR 357  
QY 598 GGTTCGGCAGGTCTCTTCCAAACATGACTTGCTGCTACGACTGACGACGCTTTTGGAC 657

DB 358 ASMNLSQAFSSLSKSLILIRIGYFELKSFNLSPHLNQLNLEVDLGNPIKIANLSMF 417  
QY 658 CAATGTCCTCGATAGAGAGAGATCGCTTCTTCGGAGAGATGTTCTTAAGCGGAAGNG 717  
DB 418 KQFKULVLDLVNKLSPSGDSSEWFGCSNARTVSEYEPVLEOLHTRDYKVASRCP 477  
QY 718 GCGGTCAATTGGCGCTGCTTCCGAGCTCGTGGCGCAAGCACTGCTTACGCTGGG 777  
DB 478 KXKESFMSVNESCYKGGTDLDSKNSIFVASSDPQHLSFLKCLNLSQNLNGSB 537  
QY 778 GTAGACATGTTACATATATAGACAGTATCGCTGTTGAGGCAAGCTTGGTCAKY 837  
DB 538 FQPLAELRYLDPSNNRLDLHSTAPEELHKLVDLSSNSHYFOSGITHMNFNKLKY 597  
QY 838 GCTTCAGGAGCGCTCTGTGTCGGCGCGCAAGTGGCGCATGATTTCTCTGCT 897  
DB 598 LQKLMNNDISSISRTWSESLRTLBERGNHLDVMBGDNRYQLFRNLKLELDI 657  
QY 898 GCATTCTGCTGTTTCTTCTCGAGCGCTTACGCGCTGCTTCGATGAGCGCGTTCCA 957  
DB 658 SKNSLSFLPSGVFDGMPNPKNLKSLAKNGLKSPWKKLOCLKNLETLDSHNLTVPER 717  
QY 958 AATCCCGGACAGTCACACTTACTTGCTACCAAGCGCTCCAAATACATGTGGAAAGCC 1017  
DB 718 LNSCSRLKNLILKNNQIRSLTYFLQDAFOLRYLDLSNKKIQIKTSPENVNLNLM 777  
QY 1018 GGGCAGCTGCGACCGAGCGTTCATCGGCTTACACGCGTGGCGCTTCTGAAG 1077  
DB 778 LLIHNRFLCTCDVAFVWVWVNTTEVTIPLATDVTCVPGAHKGSVISLDTYBELD 837  
QY 1078 GACGCTTTTCATGAGAGATATTGTGTGGCTTCGCTGCTGCTATTAATCAGCGCTTG 1137  
DB 838 TNLILFSLISVSLFLMWMNTASHLYFWDVWYIYHFCXAKIKGYORLISPDCCYDAFIY 897  
QY 1138 AATCAGACACATTAGTGCGCTCATGACTCTCGGCAATTGGCTGAACGCTTCCGG 1197  
DB 898 DTKDPATVEWVLAELVAKLEDPREKHFNLCLBERDWLPQPVLENTSOSIQLSKTVFWM 957  
QY 1198 AGGAGTCTCTCTTCAGGATAGAGAGATCTTCTGGGACACAT 1245  
DB 958 TDKYAKTENFKIAFYLSHQLMDEKVDVILLIFLEKPFQSKFLQLRK 1005

## RESULT 7

US-10-158-790-358  
; Sequence 358, Application US/10158790  
; Publication No. US2003018079A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Geriltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C448  
; CURRENT APPLICATION NUMBER: US/10/158,790  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550



```

RESULT 9
US-10-140-805-358
; Sequence 358, Application US/10140805
; Publication No. US20030207417A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C176
; CURRENT APPLICATION NUMBER: US/10/140,805
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-805-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAGGGGCGCGCATTTGACATGTGCTTATGCTGCAAACTGATGCTGAG 597
DB 298 HNSLSLGHVPPRWKINIKLQELDSLQNFLEKIGDAFLHFLPSLIQLDSLFPFELQYR 357
QY 598 GGTTCGACAGTTCCTTCCACAATGACATGCTGTACAGACGACGACGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKLIRGVFKEKSFNLSPLHNLONLEVLDTGTFIKIANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGCTTCTTCCGAGAGATGTTCCAGCCGAAAGTG 717
DB 418 KQFRKLKVIDLSVNKISPSGDSSEVGFCSNARTSVSEYPOVLEQLHYFRYDKYASCRF 477
QY 718 GCGGTCAATGGCGCTGACATTTCCGACCTGCTGTGTCGCAAAACGACTGCTTACGCTGG 777
DB 478 KNKEASFMVSNECYKYGQTLDSLKSNISIFVKSBDPQHLFLKLNLSGNLISQTLNGSE 537
QY 778 GTAGACATGTTCAATATATGAAGCAAGTATGCTGTGAGAGCAAGCTTGGTCACAT 837
DB 538 FQPLAEIRYIDFSNNRDLHSTAFEBLHLEVLDISNSHYPOSEITTMLEFKNLKV 597
QY 838 GCTTTCGAGGACGCTCTAGTGTGTCGCGCAAAATGAGGGCGAGTTCGATTTCTCTGCT 897
DB 598 LQTLMMNDNISSTSTSESLRTLEFRGNHLDVIMREGDNRYLDLFFKNLKLBLELDI 657
QY 898 GCAATTCGCTGTTGTTTCTTCTTCGACGCTTACGCGCTGTCTTTCGATGAGCGCTTCCA 957
DB 658 SKNSLSLFPBGVDFGMPNKNLSLAKNGKLSFSWKLQCLKNLETLDSLHNOQLTVPER 717
QY 958 AATCCCGGACAGTCGACACTTACTGTCTACCAAGCGCTCCCAATACATGTGAAAGCC 1017
DB 718 LSNCSRLKVLILKNQIRSLTKFYLDAPQLRLDLSSNKNIQMTQTSFPEANLNLKM 777
QY 1018 GGACAGCTGCCACGGAAGCTGTTCCATGCGGTTTACAAAGGTTGGCGCTGCTTGAAG 1077

```

```

DB 778 LLLHNRFICTDVAFFVMVWHTEVTIPIPLATDVTCVPGAHKGSQSVISLDLYTCELD 837
QY 1078 GACGCTTTCACGAGGACGACATATGCTGTGCTGCTGCGCTGTGCTTACTGACGCGCTTG 1137
DB 838 TMLILFSLISVSLFLMWMNTASHLYFMDVWYIYHFCKAKIKYQRLISPDCCYDAPIYV 897
QY 1138 AATCAGACACATTAATGAGGCGCTCATGACTCCTCGCAATTTGGCTGAACGCTTTCGGG 1197
DB 898 DTMDPAVTEWVLAELVAKLEDPREKHFNLCBERDLPQGPVLENSQISQSKTIVFM 957
QY 1198 AGGAGTCTCTTCTTCCAGGATGAGAGGATCTTTCGGACACAT 1245
DB 958 TDKYAKTENFKIAFYLSHORLMDKVDVILLIFLEKFPQSKRPLQLRK 1005

RESULT 10
US-10-140-864-358
; Sequence 358, Application US/10140864
; Publication No. US20030207419A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C184
; CURRENT APPLICATION NUMBER: US/10/140,864
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-864-358

Query Match      1.8%; Score 40.8; DB 12; Length 1049;
Best Local Similarity 5.4%; Pred. No. 0.028;
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAGGGGCGCGCATTTGACATGTGCTTATGCTGCAAACTGATGCTGAG 597
DB 298 HNSLSLGHVPPRWKINIKLQELDSLQNFLEKIGDAFLHFLPSLIQLDSLFPFELQYR 357
QY 598 GGTTCGACAGTTCCTTCCACAATGACATGCTGTACAGACGACGCTTTTGGAC 657
DB 358 ASNNLSQAFSSLSKLIRGVFKEKSFNLSPLHNLONLEVLDTGTFIKIANLSMF 417
QY 658 CAATGTTCCGATAGTGACGAGATCGCTTCTTCCGAGAGATGTTCCAGCCGAAAGTG 717
DB 418 KQFRKLKVIDLSVNKISPSGDSSEVGFCSNARTSVSEYPOVLEQLHYFRYDKYASCRF 477
QY 718 GCGGTCAATGGCGCTGACATTTCCGACCTGCTGTGTCGCAAAACGACTGCTTACGCTGG 777
DB 478 KNKEASFMVSNECYKYGQTLDSLKSNISIFVKSBDPQHLFLKLNLSGNLISQTLNGSE 537
QY 778 GTAGACATGTTCAATATATGAAGCAAGTATGCTGTGAGAGCAAGCTTGGTCACAT 837

```

Db 538 FOPLAELRYLDFSNRNLDLHSTAEBELHLEVLIDISSNSHYFOSEGITMNLFTKLV 597  
Qy 838 GCTTTCAGGAGGAGCTCTAGTCTGCGGCGGAATGGGGCGATTCGATTTCTCTGCT 897  
Db 598 LQKLMMNDNDISSSTRMSESLRLEFRGNHLDVLMREGDNRVYLQLFKNLLKLELDI 657  
Qy 898 GCATTCGCTGTTGTTTCTTCTCTCGAGCGCTTACGGCGCTGCTCGATGAGCGCTTCCCA 957  
Db 658 SKNSLSFLPSGVFDGWPMLKNLSLAKNGLSFSWKLQCLKNLETLIDLSHNDLTVPER 717  
Qy 958 AATCCCGGACAGTGCACACTTACTTGCTTACCAAGCGCTCCAAATACATGTGGAAGCC 1017  
Db 718 LSNCSRLKMLILKNNQIRSLTKYFLQDAPQLRYLIDLSNKKIQMTSPPEVNLNLKM 777  
Qy 1018 GGGCAGCTGCACCGAAGCTTCCATCGGCTTACACGGCTTGGCGCTTCTTGAAG 1077  
Db 778 LLIHNRFLCTCDVAVFWVWVWVNTTEVTIPYLATDVTCVGPAGHKGOSVSLDLYTELDL 837  
Qy 1078 GAGCGTTTCATGAGCAGATATGTTGTTGGCTTCCGCTGCTTACTTACGAGCCTTG 1137  
Db 838 TNLILFSLISVSLFLVMWMTASHLYFWDVWYIYHCKAKIKYGRLLISPDCCYDAFIY 897  
Qy 1138 AATCAGACACATTAAGTGGGCTCATGACTCTCGCAATTGGCTGAACCGTTTCGG 1197  
Db 898 DTKDPVTEWVLAELVAKLEDPREKHFNLCLEBRDWLPQGVLENLSQSIOISKTVFWM 957  
Qy 1198 AGGAGTCTCTCTCTTACGAGATAGAGAGATCTTCTGGGACACAT 1245  
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILIFLEKPFKSLQLRK 1005

## RESULT 11

US-10-140-923-358  
; Sequence 358, Application US/10140923  
; Publication No. US20030207355A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C188  
; CURRENT APPLICATION NUMBER: US/10/140, 923  
; CURRENT FILING DATE: 2002-05-07  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-923-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;

Best Local Similarity 5.4%; Pred. No. 0.028; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTTGAAGGGGCGCCATTTGACATGCTTATGCTGAACCTGATGCTGAAG 597  
Db 298 HSNLSGHVPRMVKINIKQLQELDLQNLAKELGDKAFHLFLPSLQLDLSFNFELQYVR 357

Qy 598 GGTTCGAGAGTTCCTTCCAAATCGACTTGTCTACAGCTGACAGCGTTTTCGAC 657  
Db 358 ASNNLSQAFSLSKLILIRGVFEKLSFNLSPILHNLQNEVLDTGTFIKIANLSMF 417  
Qy 658 CAATGTTCCGATAGTGAAGCAGATCGGCTTCTTCCGAGAGATGTTCTTAAGCGGAAGTG 717  
Db 418 KQFKRLKVIDLSVNKISPSGDSSEVGFCSNARTSVESYEPQVLEQHYRYDYKVASCF 477  
Qy 718 GGGTCATTGGCGCTGCGCATTTCCGACTGCTGAGTGGCAAGCAACTGCTTACGCTGG 777  
Db 478 KXKEASFMSVNESVCYVGCTLDLSKNSIFVXSPPQHLFLKCLNLISQTLNGSE 537  
Qy 778 GTAGCAGATTTACATATATAGCAAGCAAGTATCGTTGGAGGCAAGCTTGGTCAAT 837  
Db 538 FOPLAELRYLDFSNRNLDLHSTAEBELHLEVLIDISSNSHYFOSEGITMNLFTKLV 897  
Qy 838 GCTTTCAGGAGGAGCTCTTATGTCGTCGCGGCGGAATGGGGCGATGATTTCTCTGCT 897  
Db 598 LQKLMMNDNDISSSTRMSESLRLEFRGNHLDVLMREGDNRVYLQLFKNLLKLELDI 657  
Qy 898 GCATTCGCTGTTTCTTCTCTCGAGCGTTACGGCGCTGCTTTCGATGAGCGCTTCCCA 957  
Db 658 SKNSLSFLPSGVFDGWPMLKNLSLAKNGLSFSWKLQCLKNLETLIDLSHNDLTVPER 717  
Qy 958 AATCCCGGACAGTGCACACTTACTTGCTTACCAAGCGCTCCAAATACATGTGGAAGCC 1017  
Db 718 LSNCSRLKMLILKNNQIRSLTKYFLQDAPQLRYLIDLSNKKIQMTSPPEVNLNLKM 777  
Qy 1018 GGGCAGCTGCACCGAAGCTTCCATCGGCTTACACGGCTTGGCGCTTCTTGAAG 1077  
Db 778 LLIHNRFLCTCDVAVFWVWVWVNTTEVTIPYLATDVTCVGPAGHKGOSVSLDLYTELDL 837  
Qy 1078 GAGCGTTTCATGAGCAGATATGTTGTTGGCTTCCGCTGCTTACTTACGAGCCTTG 1137  
Db 838 TNLILFSLISVSLFLVMWMTASHLYFWDVWYIYHCKAKIKYGRLLISPDCCYDAFIY 897  
Qy 1138 AATCAGACACATTAAGTGGGCTCATGACTCTCGCAATTGGCTGAACCGTTTCGG 1197  
Db 898 DTKDPVTEWVLAELVAKLEDPREKHFNLCLEBRDWLPQGVLENLSQSIOISKTVFWM 957  
Qy 1198 AGGAGTCTCTCTCTTACGAGATAGAGAGATCTTCTGGGACACAT 1245  
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILIFLEKPFKSLQLRK 1005

## RESULT 12

US-10-141-756-358  
; Sequence 358, Application US/10141756  
; Publication No. US20030207359A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C200  
; CURRENT APPLICATION NUMBER: US/10/141, 756  
; CURRENT FILING DATE: 2002-05-08

; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 358  
 ; LENGTH: 1049  
 ; TYPE: PR1  
 ; ORGANISM: Homo Sapien  
 US-10-141-756-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;  
 Best Local Similarity 5.4%; Pred. No. 0.028;  
 Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCCGCGCATTTGACATGTAAGTGTGCTGAACCTGATGCTGAAG 597  
 DB 298 HNSLSLQHPVPRWKNINKLOEIDLSQFLAKEIDAKFLHFLPSLIQLDLSFNELOVR 357  
 QY 598 GGTTCGGAGGCTTCCTTCCAAACATGCACTGCTTACAGATGACAGCCGTTTTTGAC 657  
 DB 358 ASNMLSQAFSSLKSLKILIRIGVFEKLSFNLSPILNLEVLDTGNFIKIANLSMF 417  
 QY 658 CAATGTTCCATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCGGAAGTG 717  
 DB 418 KQFRILKVIDLSVNKISPSGDSSEVGFCSNARTSVSEYEPQVLEQLHYFRYDKYASCRF 477  
 QY 718 GCGGTCATGCGCGTGGCATTTCCGAGCTGCTGAGTGGCAACGACTGCTTCACTGCGG 777  
 DB 478 KKEASFMVSNECYKRGOTLDLSKNSIFVKSDDPHLSFLKLNLSGNLSIQTUNGSE 537  
 QY 778 GTAGACAGATTTACATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGCTCAT 837  
 DB 538 FQPLAELRYLDFSNRDLHSTAFELHKLVDLSSNSHYQSEGITMLNFTKLV 597  
 QY 838 GCTTTCAGGAGCGCTCTAGTGTGCTGCGGAAATGGGGCGAGCTTCCCTGCT 897  
 DB 598 LQKLMMNDNDISSSTRMESELRTLEFRGNHLDVLMREGDNRVQLDFNLLKLELDI 657  
 QY 898 GCATTCGCTGTTGTTCTTCTCGAGAGCTTACGCGCTGCTGATGAGCGCGTCCA 957  
 DB 658 SKNSLSLPSGVDPGMPNLKNLSLAKNGLSFSWKLOCLKNLETLIDLSHNLTVPER 717  
 QY 958 AATCCCGGACAGTGCAGACTTACTGTTGTTACCAAGCGCTCAATCATGTGGAAGCC 1017  
 DB 718 LSNCSRLKMLILKNNQIRSLTKFLQDAFQRLYLDLSNKNIQMIOKTSPEVANLNM 777  
 QY 1018 GGGCAGCTGCCAGGAGCTGTTCCATCGCGTTTACACGCTGGCGCTTCTTAAG 1077  
 DB 778 LLLHNRFLCTGDAVFWFVMMVNHTEVTIPLYATDVTCVGRGAKGQSVISLDLYTCBLD 837  
 QY 1078 GACGCTTTTCATGAGCAGATATGTTGCTGCTGCTGCTGCTTACTACAGCGCTTG 1137  
 DB 838 TNLILFSLISVSLFLMVMNTASHLYFWDWYIYHFCAKIKYQRLISPDCCYDAIIV 897  
 QY 1138 AATTCAGACACATTAAGTGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTGGG 1197  
 DB 898 DTDPATVTEWVLAELVAKLEDPREKFNLCLEERDMLPGQVLENLQSISQISKYIVFM 957  
 QY 1198 AGGAGTCTTCTTCTTCAAGGATAGAGATCTTCTTGGGCAACAT 1245  
 DB 958 TDKYAKTENFKIAFYLSHORLMBEKVDVILLIFLEKPFQSKFLQLRK 1005

RESULT 13  
 US-10-141-759-358  
 ; Sequence 358, Application US/10141759  
 ; Publication No. US20030207361A1  
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Defoige, Laura  
 ; APPLICANT: Deenoys, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerlicsen, Mary E.

; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowsky, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria A.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P3330R1C197  
 ; CURRENT APPLICATION NUMBER: US/10/141,759  
 ; PRIOR APPLICATION: 2002-05-08  
 ; PRIOR APPLICATION removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 358  
 ; LENGTH: 1049  
 ; TYPE: PR1  
 ; ORGANISM: Homo Sapien  
 US-10-141-759-358

Query Match 1.8%; Score 40.8; DB 12; Length 1049;  
 Best Local Similarity 5.4%; Pred. No. 0.028;  
 Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

QY 538 ACCCTTGAAGGGCCGCGCATTTGACATGTAAGTGTGCTGAACCTGATGCTGAAG 597  
 DB 298 HNSLSLQHPVPRWKNINKLOEIDLSQFLAKEIDAKFLHFLPSLIQLDLSFNELOVR 357  
 QY 598 GGTTCGGAGGCTTCCTTCCAAACATGCACTGCTTACAGATGACAGCCGTTTTTGAC 657  
 DB 358 ASNMLSQAFSSLKSLKILIRIGVFEKLSFNLSPILNLEVLDTGNFIKIANLSMF 417  
 QY 658 CAATGTTCCATAGTGAAGCATGCGCTTCTTCCGAGAGATGTTCTTAAGCGGAAGTG 717  
 DB 418 KQFRILKVIDLSVNKISPSGDSSEVGFCSNARTSVSEYEPQVLEQLHYFRYDKYASCRF 477  
 QY 718 GCGGTCATGCGCGTGGCATTTCCGAGCTGCTGAGTGGCAACGACTGCTTCACTGCGG 777  
 DB 478 KKEASFMVSNECYKRGOTLDLSKNSIFVKSDDPHLSFLKLNLSGNLSIQTUNGSE 537  
 QY 778 GTAGACAGATTTACATATATGAAGCAAGTATGCTGTTGAGGCAAGCTTTGCTCAT 837  
 DB 538 FQPLAELRYLDFSNRDLHSTAFELHKLVDLSSNSHYQSEGITMLNFTKLV 597  
 QY 838 GCTTTCAGGAGCGCTCTAGTGTGCTGCGGAAATGGGGCGAGCTTCCCTGCT 897  
 DB 598 LQKLMMNDNDISSSTRMESELRTLEFRGNHLDVLMREGDNRVQLDFNLLKLELDI 657  
 QY 898 GCATTCGCTGTTGTTCTTCTCGAGAGCTTACGCGCTGCTGATGAGCGCGTCCA 957  
 DB 658 SKNSLSLPSGVDPGMPNLKNLSLAKNGLSFSWKLOCLKNLETLIDLSHNLTVPER 717  
 QY 958 AATCCCGGACAGTGCAGACTTACTGTTGTTACCAAGCGCTCAATCATGTGGAAGCC 1017  
 DB 718 LSNCSRLKMLILKNNQIRSLTKFLQDAFQRLYLDLSNKNIQMIOKTSPEVANLNM 777  
 QY 1018 GGGCAGCTGCCAGGAGCTGTTCCATCGCGTTTACACGCTGGCGCTTCTTAAG 1077  
 DB 778 LLLHNRFLCTGDAVFWFVMMVNHTEVTIPLYATDVTCVGRGAKGQSVISLDLYTCBLD 837  
 QY 1078 GACGCTTTTCATGAGCAGATATGTTGCTGCTGCTGCTGCTTACTACAGCGCTTG 1137  
 DB 838 TNLILFSLISVSLFLMVMNTASHLYFWDWYIYHFCAKIKYQRLISPDCCYDAIIV 897  
 QY 1138 AATTCAGACACATTAAGTGGCTCATGACTCTGCGAAATTTGGCTGAACGTTTGGG 1197  
 DB 898 DTDPATVTEWVLAELVAKLEDPREKFNLCLEERDMLPGQVLENLQSISQISKYIVFM 957  
 QY 1198 AGGAGTCTTCTTCTTCAAGGATAGAGATCTTCTTGGGCAACAT 1245

Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILILEKPEFKSKFLQLRK 1005

# RESULT 14

US-10-123-155-358

; Sequence 358, Application US/10123155  
; Publication No. US20030068794A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C30  
; CURRENT APPLICATION NUMBER: US/10/123,155  
; CURRENT FILING DATE: 2002-04-15  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
; US-10-123-155-358

Query Match 1.8%; Score 40.8; DB 14; Length 1049;  
Best Local Similarity 5.4%; Pred. No. 0.028;  
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTGAAGGGGCGCGCATTTGACATGATGATGCGAAGCTGATGCTGAAG 597  
Db 298 HNSLQHVPRPFWKINIKLOEIDLSQFLAKEIDAKFLHFLPSLIDLSFNFELQVYR 357  
Qy 598 GGTTCGGCAGGATTCCTTCCACATGACATGCTTACGACGACGCGTTTGTGAC 657  
Db 358 ASNMLSQAFSLSKLIRIGVYFKELKSPNLSPLHNLQNLVLDLGTNFIKIANLSMF 417  
Qy 658 CAATGTTCCGATGATGACGATGCGCTTCTTCCGAGATGTTCTTAAGCGGAAGTG 717  
Db 418 KQFRRLKVIDLSVVKISPSGDSSEVFCNSARTSVESYEPQVLEQLHYFYDYKXARSCR 477  
Qy 718 GCGGTCATGGCGGTCGATTTCCGACGCTGCTGTCGCAAGCAAGCTTACGCTGAG 777  
Db 478 KKEEASFMVNSCYVYGGTDLDSKNSIFVKSDFPHLSFKLNLNSNLISQTLNGSE 537  
Qy 778 GTAGACAGATGTTACATATATGAAGCAAGTATGCTGTTGAGCAAGCTTGTGTCACAT 837  
Db 538 FQPLAEARLYDFSNRDLHSTAFELHKLLEVLDISNSHYQSEBGTIMLNFTRKLV 597  
Qy 838 GCTTTCGAGGACGCTCTAGTGTGTCGCGAAGTGGGGCGATGCGATTTCTCTGCT 897  
Db 598 LQKLMNDNDISSSTRTMESLRTLEFRGNHLDVLMREGDNRYLQLFKNLKLEBLDI 657  
Qy 898 GCATTCGCTGTTGTTTCTTCCTCGACGCTTACGCGCTGCTTCGATGAGCGGCTTCCA 957  
Db 658 SKNSLSLFBGVDPMPNKLKNSLAKGKLSFKWKLQCLKULETLDLSHNDLTVPER 717  
Qy 958 AATCCCGGACAGTGCATGCTACTGCTTACCAAGCGCTCCATATCATATGTAAGGCC 1017  
Db 718 LSNCSRLKVLILKNNQIRSLTKYFLQDAFQRLYLDLSSNKLQIMQKTSPEVIANLKLK 777

Qy 1018 GGGCAGGTGCGACGAGCTTCCATCGGCTTTACAGCGTTGGGCGCTTCTGAAG 1077  
Db 778 LLLHNRPLCTCAVFWVWVNHTEVITPYLATDVTCVGRANKGQSVSLDYTBELD 837  
Qy 1078 GACGTTTTCATGAGAGATATTGTTGGCTTCGCTGCTGCTTATTACTAGCGCTTG 1137  
Db 838 TNLILPSLSISVSLFLMVMWMTASHLYFWDVWYIYHCKKAKIKGYQLISPDCCYDAFIV 897  
Qy 1138 AATTCAGACACATTAGTGCGGCTCATGACTCTCGCAATTGGCTGAACGCTTTCGG 1197  
Db 898 DTDPAVTEWVLAELVAKLEDPREKHNLCLERDMLPGQPVLENLSOSIQLSKKTIVFW 957  
Qy 1198 AGGAGCTCTCTTCGAGGATAGAGAGATCTTCTGGGACACAT 1245  
Db 958 TDKYAKTENFKIAFYLSHORLMDKVDVILILEKPEFKSKFLQLRK 1005

# RESULT 15

US-10-146-731-358

; Sequence 358, Application US/10146731  
; Publication No. US2003012952A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C323  
; CURRENT APPLICATION NUMBER: US/10/146,731  
; CURRENT FILING DATE: 2002-05-15  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
; US-10-146-731-358

Query Match 1.8%; Score 40.8; DB 15; Length 1049;  
Best Local Similarity 5.4%; Pred. No. 0.028;  
Matches 38; Conservative 197; Mismatches 473; Indels 0; Gaps 0;

Qy 538 ACCCTGAAGGGGCGCGCATTTGACATGATGATGCGAAGCTGATGCTGAAG 597  
Db 298 HNSLQHVPRPFWKINIKLOEIDLSQFLAKEIDAKFLHFLPSLIDLSFNFELQVYR 357  
Qy 598 GGTTCGGCAGGATTCCTTCCACATGACATGCTTACGACGACGCGTTTGTGAC 657  
Db 358 ASNMLSQAFSLSKLIRIGVYFKELKSPNLSPLHNLQNLVLDLGTNFIKIANLSMF 417  
Qy 658 CAATGTTCCGATGATGACGATGCGCTTCTTCCGAGATGTTCTTAAGCGGAAGTG 717  
Db 418 KQFRRLKVIDLSVVKISPSGDSSEVFCNSARTSVESYEPQVLEQLHYFYDYKXARSCR 477  
Qy 718 GCGGTCATGGCGGTCGATTTCCGACGCTGCTGTCGCAAGCAAGCTTACGCTGAG 777  
Db 478 KKEEASFMVNSCYVYGGTDLDSKNSIFVKSDFPHLSFKLNLNSNLISQTLNGSE 537

```
OY 778 GTAGACATGTACAAATATATGAAGCAAGTATCGTTGTGAGGCAAGCTTTGGTCACAT 837
Db 538 FQPLAEIRYIDPNNRDLDLHSTA F E L H L E V L D I S N S H Y P Q S E G I T M L N F T K N L K V 597
OY 838 GCTTTGAGGAGCGCTCTCTAGTGTGCGCGGAAATGGGGCGATGCGATTCTCTGCT 897
Db 598 LQRLMNDNDISSSTSTMESESRLTEFRGNHLDVLMREGDNRVYLQLFKNLLKLELDI 657
OY 898 GCATTCTGCTGTTTTCTTCTCGAGCGTTACGGCGCTGTCTCGATGAGCGCGTCCCA 957
Db 658 SKNSLSFLPSGVFDGMPNLIKNLSLAKNGLKSPSWKLOCLKNLETLDLSHNQLTVPER 717
OY 958 AATCCGCGACAGTCGACACTTACTTGTCTACCAAGCGCTCCCAATACATGAGGAAAGCC 1017
Db 718 LSNCSRSLLKRLILKNNDIRSLTKYFLQDAFQLRYLDLSNKKIQMIQKTSFPENVLNNLKM 777
OY 1018 GGGCAGCTGCCACGAGCTGTTCCATCGCGTTTACAAAGGTTGGCGGTGCGTTCTGAAG 1077
Db 778 LLLHNRFLCTCDVAVMFVMMVNNHTEVTIPYLATDVTCVPGAHKQGSVISLDLYTCEL DL 837
OY 1078 GACGGTTTTCATGAGCGAGATATTGTGTGGCTTCGCTCGCTATTA CTCAAGCCTTG 1137
Db 838 TNLILPSLSISVSLFLMVMWMTASHLYFWDVWYIYHFKAKIKGYQRLISPDCCYDAFIY 897
OY 1138 AATCAGAGACACATTAAGTGGGCTCATGACTCTCGCAATTGGCTGAACGTTTCGGG 1197
Db 898 DTGDPATVEWVLELVAKLEDPREKHFNLLERDMLPGQPVLENLSQSIOLSKTVFVM 957
OY 1198 AGGAGTCTTCTCTTCAGGAGATAGAGAGATCTTTCTGSGCACACAT 1245
Db 958 TDKYAKTENFKIAFYLSHQRLMDKDVILILIEKPFQKSKFLQLRK 1005
```

Search completed: November 23, 2003, 21:22:45  
Job time : 709.748 secs



Dh 5929 CGAAGACCTTCTAGGGGAGAGGATTACTCAATCTCAACGAGTGCAGCTGGGTTA 5988  
Qy 181 GCTTGCAGAAAGCTGGCCGATGGTCCGTTCCCGAGATCTCAGCTGGTGGAAAGTATGCA 240  
Db 5989 GCTTGCAGAAAGCTGGCCGATGGTCCCGAGATCTCAGCTGGTGGAAAGTATGCA 6048  
Qy 241 GTTCTCTCCGCTTATATCTATTTATTTGGCAAGAAATTTGGGGCGGATCTTGAATCGAAA 300  
Db 6049 GTTCTCTCCGCTTATATCTATTTATTTGGCAAGAAATTTGGGGCGGATCTTGAATCGAAA 6108  
Qy 301 CCTTGGCGGGGCAACAGTGAAGTGTCTGTTGCCATGCACTTGAACACCATTTGCATG 360  
Db 6109 CCTTGGCGGGGCAACAGTGAAGTGTCTGTTGCCATGCACTTGAACACCATTTGCATG 6168  
Qy 361 GATTTTCCGCAACCACTTAATCCAAAGCCCTGTTTGGTGAACGGTTAAAGATGCA 420  
Db 6169 GATTTTCCGCAACCACTTAATCCAAAGCCCTGTTTGGTGAACGGTTAAAGATGCA 6228  
Qy 421 CCGATTGATCTTAGTCACTTTCGTGGCCATTCAATCTTAAGACTGCGGCTTTGCAAC 480  
Db 6229 CCGATTGATCTTAGTCACTTTCGTGGCCATTCAATCTTAAGACTGCGGCTTTGCAAC 6288  
Qy 481 CTGCAATGCGCGTGTACAGAAATGGCAGATGAATGCGTTACCGGTTTACCATACC 540  
Db 6289 CTGCAATGCGCGTGTACAGAAATGGCAGATGAATGCGTTACCGGTTTACCATACC 6348  
Qy 541 CTTGAAAGGGGCGGTGCCATTTGACATAGTACCTTAATAGTGGAAACCTGAATGCTGAAGGT 600  
Db 6349 CTTGAAAGGGGCGGTGCCATTTGACATAGTACCTTAATAGTGGAAACCTGAATGCTGAAGGT 6408  
Qy 601 TCGGACAGGTTCTTTCCCAACAATCGACTTCTCAACACTGACAGACCGTTTGTGACCA 660  
Db 6409 TCGGACAGGTTCTTTCCCAACAATCGACTTCTCAACACTGACAGACCGTTTGTGACCA 6468  
Qy 661 TGTTCGATATGAGACGAGATCGGCTTCTTCCGAGAGATTTCTTAAGCCGAAATGGCG 720  
Db 6469 TGTTCGATATGAGACGAGATCGGCTTCTTCCGAGAGATTTCTTAAGCCGAAATGGCG 6528  
Qy 721 GTCAATGGCCCTGGCATTTCCGACTGTGTGTGGCAACGAACTGCTTATGCTGGGGTA 780  
Db 6529 GTCAATGGCCCTGGCATTTCCGACTGTGTGTGGCAACGAACTGCTTATGCTGGGGTA 6588  
Qy 781 GACGATGTCAATATATGAGCAAGTATCGTGTGGAGCAAGCTTGTATGCTCAATGCT 840  
Db 6589 GACGATGTCAATATATGAGCAAGTATCGTGTGGAGCAAGCTTGTATGCTCAATGCT 6648  
Qy 841 TTCAGGACGCTCTAGTGTGTGGCCGAAATGGGGCGATGCGATTTCTCTGCTGCA 900  
Db 6649 TTCAGGACGCTCTAGTGTGTGGCCGAAATGGGGCGATGCGATTTCTCTGCTGCA 6708  
Qy 901 TTCTGCTTTTCTTCTCTCGAGCCGTTACGGCTGTCTTGCATGAGCCGTTCCCAAT 960  
Db 6709 TTCTGCTTTTCTTCTCTCGAGCCGTTACGGCTGTCTTGCATGAGCCGTTCCCAAT 6768  
Qy 961 CCCGGCAGTGCAGACTTACTTGTGTCTACCAAGGCGTCCCAATCATGTGAAAAGCCGG 1020  
Db 6769 CCCGGCAGTGCAGACTTACTTGTGTCTACCAAGGCGTCCCAATCATGTGAAAAGCCGG 6828  
Qy 1021 CAGCTGCAACGAGCTGTTCCATCGCGTTTACACGCTTGGCGTGGCTTTTGAAGAC 1080  
Db 6829 CAGCTGCAACGAGCTGTTCCATCGCGTTTACACGCTTGGCGTGGCTTTTGAAGAC 6888  
Qy 1081 GGTTTTATAGAGAGATATTTGTGTGGCTTCCGCTGTCTTACTCAGGCTTTGAAA 1140  
Db 6889 GGTTTTATAGAGAGATATTTGTGTGGCTTCCGCTGTCTTACTCAGGCTTTGAAA 6948  
Qy 1141 TCAGACACATTAAGTGGGCTCATGACTCCGCAAAATTTGGCTGAACCGTTTGGAGG 1200  
Db 6949 TCAGACACATTAAGTGGGCTCATGACTCCGCAAAATTTGGCTGAACCGTTTGGAGG 7008  
Qy 1201 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTTGGGCAACATCTCTGTGTGT-AA 1259  
Db 7009 GAGTCTTCTCTTCAAGGATAGAGAGATCTTCTTGGGCAACATCTCTGTGTGTGAA 7068

Qy 1260 ACATGAGTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGAATAGGATCTGGCGG 1319  
Db 7069 ACATGAGTTTCTCATGATTTGGGACCTATTCAAGCTAATGGGAATAGGATCTGGC-GG 7127  
Qy 1320 GTTTGGTCCAGTTTGTGAAAGCGGGTTTATTTAGATCTCTCCGTTGGTCAATCAACGATA 1379  
Db 7128 GTTTGGTCCAGTTTGTGAAAGCGGGTTTATTTAGATCTCTCCGTTGGTCAATCAACGATA 7287  
Qy 1380 TGAAGAAATCAGCGGATGTCCCTGAAGAAATCTCAGAACTTCCACGTGGAATCCGATC 1439  
Db 7188 TGAAGAAATCAGCGGATGTCCCTGAAGAAATCTCAGAACTTCCACGTGGAATCCGATC 7247  
Qy 1440 TGAAGTTTAAAGGTTGTCTGTGAGCCAGCATATGTCATGTTCAAGTCAAGCGGAT 1499  
Db 7248 TGAAGTTTAAAGGTTGTGTGTGAGCCAGCATATGTCATGTTCAAGTCAAGCGGAT 7307  
Qy 1500 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTTATGA 1559  
Db 7308 TCAGAGGAAAGACAAATAAAGATTAAGGCTTAAGAGCGGATATCTGAACCTTTATGA 7267  
Qy 1560 TAAAGTGTGTCACTCTGACTCCGAAATATCCAACTCAGGCAATTCCTGCAATGCCA 1619  
Db 7368 TAAAGTGTGTCACTCTGACTCCGAAATATCCAACTCAGGCAATTCCTGCAATGCCA 7427  
Qy 1620 TACCAATTTTTTCAAGCAACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 1679  
Db 7428 TACCAATTTTTTCAAGCAACAGTGAACCAAGCGGTTGATTAACAGCCATATGACAGATC 7487  
Qy 1680 GTCAAAACTCTTCTGATGATGATGAAGAAATTTCTGTTATAGACATATCTCCGCTTTS 1739  
Db 7488 GTCAAAACTCTTCTGATGATGATGAAGAAATTTCTGTTATAGACATATCTCCGCTTTS 7547  
Qy 1740 TGTCTCATGAGAGGATGCGCAAAAGCAGTGTATTCCTGGAATATGAGTGCAGAGATCC 1799  
Db 7548 TGTCTCATGAGAGGATGCGCAAAAGCAGTGTATTCCTGGAATATGAGTGCAGAGATCC 7607  
Qy 1800 GAATGTAAAGGCTAGTGTCTCATATACATGAGAGAGAGATCCCAAGCTGTT 1859  
Db 7608 GAATGTAAAGGCTAGTGTCTCATATACATGAGAGAGAGATCCCAAGCTGTT 7667  
Qy 1860 GGGGCTCCCGACAAAGAGGATTAATGTCTGCTGCGGAGCGCAATTTGAGATCTTT 1919  
Db 7668 GGGGCTCCCGACAAAGAGGATTAATGTCTGCTGCGGAGCGCAATTTGAGATCTTT 7727  
Qy 1920 CCCGGCTTTGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTCA 1979  
Db 7728 CCCGGCTTTGCCAGACCTATTTCTGCTGCTGCTGATTAAGACCAAAATGTTATTCA 7787  
Qy 1980 ACAATGATGCTTACAGAGAGATGCGGGGAGGCTTTCAAACTCAACCGGCGTGTGA 2039  
Db 7788 ACAATGATGCTTACAGAGAGATGCGGGGAGGCTTTCAAACTCAACCGGCGTGTGA 7847  
Qy 2040 GGAATTTTATTTGAAGAACTTTTCTTCAAGCACTGACACGCGTAAATGATCCGAGT 2099  
Db 7848 GGAATTTTATTTGAAGAACTTTTCTTCAAGCACTGACACGCGTAAATGATCCGAGT 7907  
Qy 2100 TTACTTGGCGGTTGCAATGTTCTTCAAGTGTGATGGTGTGAGGCTTATTCAGAC 2159  
Db 7908 TTACTTGGCGGTTGCAATGTTCTTCAAGTGTGATGGTGTGAGGCTTATTCAGAC 7967  
Qy 2160 CCGGTGTAACGGCGTGTGCAATTAATCAATTTGAGAGGCAATTTGGCAAGGCA 2219  
Db 7968 CCGGTGTAACGGCGTGTGCAATTAATCAATTTGAGAGGCAATTTGGCAAGGCA 8027  
Qy 2220 TCTCTGAAACACTTGTGAGAGATATACTACCCAGTGAATAATTAG 2268  
Db 8028 TCTCTGAAACACTTGTGAGAGATATACTACCCAGTGAATAATTAG 8076

RESULT 2  
US-09-073-587-2  
; Sequence 2, Application US/09073587

```

? Patent No.6268552
? GENERAL INFORMATION:
? APPLICANT: U.S. YI
? TITLE OF INVENTION: Transgenic Seedless Plants
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
? STREET: 5370 Manhattan Circle, Suite 201
? CITY: Boulder
? STATE: Colorado
? COUNTRY: US
? ZIP: 80303
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/073,587
? FILING DATE: 06-MAY-1998
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/045,725
? FILING DATE: 06-MAY-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Fether, Donna M.
? REGISTRATION NUMBER: 33,878
? REFERENCE/DOCKET NUMBER: 4-97
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (303) 499-8080
? TELEFAX: (303) 499-8089
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2211 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 4..2205
? JS-09-073-587-2

```

```

Db      1440 GAAGGAAAGACAAAATAAGATTAAGCTTAAGAGGGGATATCTGAATCTTTATGATA 1499
Qy      1563 GGTGGTGTGACATCTGGACTGCGAAATATCCAACTCAGGCAATTCCTGATCGATATAC 1622
Db      1500 GGTGGTGTGACATCTGGACTGCGAAATATCCAACTCAGGCAATTCCTGATCGATATAC 1559
Qy      1623 CAATATTTTCAAGGACCACTGAACCAAGGGTGTGATTAAGCCATATGACAGATCTGC 1682
Db      1560 CAATATTTTCAAGGACCACTGAACCAAGGGTGTGATTAAGCCATATGACAGATCTGC 1619
Qy      1683 AAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1742
Db      1620 AAAACTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
Qy      1743 CCTCATGAGAGGAGATGCGAAAGAGTGTATGCTGAGATGATGATGATGATGATGATGATGAT 1802
Db      1680 CCTCATGAGAGGAGATGCGAAAGAGTGTATGCTGAGATGATGATGATGATGATGATGATGAT 1739
Qy      1803 TGGTAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1862
Db      1740 TGGTAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799
Qy      1863 GGTCCCGACAAAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1922
Db      1800 GGTCCCGACAAAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1859
Qy      1923 GGGGTTTCCGACCACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1982
Db      1860 GGGGTTTCCGACCACTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1919
Qy      1983 TGAATGGCTTACAGACAGAAATGCGGGGAGCTTTAACTCAACCGGGGTGTGAGGA 2042
Db      1920 TGAATGGCTTACAGACAGAAATGCGGGGAGCTTTAACTCAACCGGGGTGTGAGGA 1979
Qy      2043 TTTTATTTCTGAAGACTTTTCTTTCAAGCACTGACACCGGCTATGATATCCGAGTTTA 2102
Db      1980 TTTTATTTCTGAAGACTTTTCTTTCAAGCACTGACACCGGCTATGATATCCGAGTTTA 2039
Qy      2103 CTGGCGGGGTGACAGTGTCTTTCACAGGTGATGGGTGAGGGTGTATTCAGACCGC 2162
Db      2040 CTGGCGGGGTGACAGTGTCTTTCACAGGTGATGGGTGAGGGTGTATTCAGACCGC 2099
Qy      2163 GTGTAAAGCGGTCTGTGCAATTAATCCCAATGTGTGAGGCAATTTGGCAAGGGCAATCC 2222
Db      2100 GTGTAAAGCGGTCTGTGCAATTAATCCCAATGTGTGAGGCAATTTGGCAAGGGCAATCC 2159
Qy      2223 TCTGAAACACTCTTGAAGAGATTAATACCCGACAGAGAAATTAAG 2268
Db      2160 TCTGAAACACTCTTGAAGAGATTAATACCCGACAGAGAAATTAAG 2205

```

## RESULT 3

```

US-08-008-216-8
; Sequence 8, Application US/08008216
; Patent No. 5366887
; GENERAL INFORMATION:
; APPLICANT: Slightom, Jerry L.
; APPLICANT: Tepfer, David A.
; TITLE OF INVENTION: R1 T-DNA Promoters
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY, CARY, AMES & FRYE
; STREET: 401 B Street, Suite 1700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-4297
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,216
; FILING DATE: 25-JAN-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/725,368
; FILING DATE: 22-APR-1985
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhorst, Marnie W.
; REGISTRATION NUMBER: 36,740
; REFERENCE/DOCKET NUMBER: P1020US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 699-2700
; TELEFAX: (619) 236-1048
; INFORMATION FOR SEQ. ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Agrobacterium rhizogenes
; STRAIN: Strain A4
; IMMEDIATE SOURCE:
; LIBRARY: Convulvulus arvensis plant cells
; CLONE: Clone 7
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..2279
; OTHER INFORMATION: /label= ORF8
; OTHER INFORMATION: /note="Sequence ORF8 corresponds to bases 6609
; OTHER INFORMATION: through 8888 of Seq. ID No. 5366887 19. "
US-08-008-216-8

```

```

Query Match      6.0%; Score 136.4; DB 1; Length 2279;
Best Local Similarity 45.9%; Pred. No. 8.7e-37;
Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;

```

```

Qy      623 TCGACTTGCTCTACGACTGACGACCGCTTTTTCACCAATGTTCCGATGATGACGAGATCG 682
Db      665 TTGACCTGCGGTACGACTATACAGTGTCTTCTCCAGAGAGTGCATGCCATGATGATGAG 724
Qy      683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGCTATTTGGCGGCTGATTTCCG 742
Db      725 GGCATTATCCGAGAGCGACCACTACCTAGTGTTCATCGTAGAGAGGGGCTGTCTG 784
Qy      743 GACTCGTGTGGCAAGCAAGATGCTTCATGCTGGGGTAGACGATGTTACATATATGAG 802
Db      785 GCCTTGTTGCTGCACAGAACTACTTGGGCGTGGCGTCAAGAAATCACTTTTGCATA 844
Qy      803 CAAGTATGCTGTGAGGCA--GCTTGTGATGATGCTTTTCAGGAGAGCTCTAGTG 859
Db      845 CCGTTATGAGATCCGATGTTTGGGGCAATCGCGCATGCCAAGCGGAGCTCACAGG 904
Qy      860 TCGTGGCCGAAGAGGGGGGAGATGCAATTCCTCTGCTGATCTGCTGTTTCTTCC 919
Db      905 CTTGACGTGCTTGGGTGATGATCTTTCTCCGCAACCAACTTTGCTCTGATCTATC 964
Qy      920 TCGAGGTTAAGGCGCTGCTTTCGATGAGGCGGTTCCCAATCCCGGACAGTGCACATT 979
Db      965 TGAATAAGTTTAATATTCGTCAGGCTTCTTCTTCTGTGCGGGAAGACACAGCAG 1024
Qy      980 ACTTGTCTACAGAGGCTCAATATGATGAGAAAGCGGAGCTGCCACCGAAGCTGT 1039
Db      1025 CACTATATTTCCCGCAGAAAGCTACGATGGCAGCGGGGCAAGCTCCCGGGGATAT 1084
Qy      1040 TCCATGCGCTTACACAGCTTGGCGGTGCTTCTTGAAGAGCGGTTTTCATGACGAGATA 1099
Db      1085 TTCAGCGGGTACATGTGAGTGAAGACACTACTCTTACCAAGGATGGAACGAAATGCA 1144
Qy      1100 TTGTGTTGGCTTGCGCTGCTATTAATCAAGGCTTGAATACAGACATTAAGTGGG 1159

```

|    |      |   |      |
|----|------|---|------|
| Db | 1145 | GGAGCTGATGCTCCGATGGAIVNCTTTTCATSTGAAAGAGCGTGTGTATGAAG             | 1204 |
| Qy | 1160 | CTCATGACTCTTGCGCAAAATTTGGCTGAACCGTTTGGGAGGGAGTCTTCTCTTCAGGGA      | 1219 |
| Db | 1205 | CCTGAGAAGCAGCGGAGCTTTGGCTCCGAGAGTTGGAAAAATTCATCTTTCATGCGCTTT      | 1264 |
| Qy | 1220 | TAGAGAGGATCTTCTCTGGGCAACAATCTCCCTGGTGT-----AAACATGAGATTTTCTC      | 1275 |
| Db | 1265 | TGCTCGAGATCTTCAGCTGTGGTAATTTGAGTCCCTGTGGCAGAGGCATGGCAACACCCC      | 1324 |
| Qy | 1276 | ATGATTTGGACCTATTCAGGCTAATGGAAATAGATCTGGCGGGTTTGTCCAGTTT           | 1335 |
| Db | 1325 | ATGATTTTCGAGGCTTTCCGGGAAATACGAGTTTGGGATACGCGCCGAGTTTCGTCC-TATTAAC | 1383 |
| Qy | 1336 | GAAGCGGCTTATTTGAGATCTCCGTTGGTCATCAACGATATGAAAGAAATACAGCG          | 1395 |
| Db | 1384 | AACCTGTTGTTTTCACGATCCTTGACCTGAGATTAATCAATGGCTACGAGAGACCAAGAT      | 1443 |
| Qy | 1396 | ATGTGCCCTGAAAGAAATCTCAGAACTTTCACGTCGGATCGCATGGAATGGTTAAACGCT      | 1455 |
| Db | 1444 | CTTCTCTATTTGGTGGGGTTCAACTTTTGACGCTCGATGGCCATTGAATATTCACAGAA       | 1503 |
| Qy | 1456 | GTCCTGTGAGCCACGCGCATATGCCATTTCAAGTCAGGGCGATTCAGAGAAAAGACA         | 1515 |
| Db | 1504 | AGCCATCGAAAGACGACGACTGTGTTGATCCCGCTGTGAATAAGCCMAAGAGGCGCGG        | 1563 |
| Qy | 1516 | AAATTAAGATPAAGGCTTAAGAGGGGATATCTGAACCTTATPAGATGGTGGTACA           | 1575 |
| Db | 1564 | AGATTTGAAGGTATCTGTAAACAGGTATTCGGGTGTTTGTGACAGAGTATCATTTGGC        | 1623 |
| Qy | 1576 | TCTGGACTCGCAAAATATCCAACTGACGAGTTCCTGACATGCGATACCAATATTTTTCAG      | 1635 |
| Db | 1624 | GGCAGTCTGAGCGCGCTACAGTTGATPAACAGACTGGCGGGGATGAGACTTCTTCAGC        | 1683 |
| Qy | 1636 | GCACAGTGAACCAAGCGGTGATPAACAGCCATATGAPAGAGATGTCAAAATCTTCTCTG       | 1695 |
| Db | 1684 | TACAAATATCGAACCCGCGTCGAAACCTCGTCTGCGCTGTCAATTACGACACTCTTCAATG     | 1743 |
| Qy | 1696 | ATGACTGAACCAAAATTTCTGGTTAGACCATATTCCTCCGCTTGTGTCTCATGAGACGG       | 1755 |
| Db | 1744 | GTCAGAAAGCAAAAGTTTGGGTTAATCTCGGGCANTCCAGCAGATATGAGACCGATGGG       | 1803 |
| Qy | 1756 | ATCGCAAAAGCAGTGTATTTGCCCTGAGCTATATGATGACAGGATCCGAATGTGTAAGGCTTA   | 1815 |
| Db | 1804 | CTTGTCCGTGAGCTGTGTGTGATTCATTCAGATCG-----CCAGCTGGAGAGGGCCTT        | 1857 |
| Qy | 1816 | GTCGTCATCAGTTATPAACATGGAGAGACATCCCAACAAAGCTTGGGCGGTCCCGACAA       | 1875 |
| Db | 1858 | GTCGTTTTTACTATGCTTGTGGATGACATATATCCCGGCGGATCGACATATGACAAAG        | 1917 |
| Qy | 1876 | AAAGAGCGATTATATCTGCTGCGGAGCGCAATTTGAGATCTTTCCGCGCTTTGCCAG         | 1935 |
| Db | 1918 | AAGGAGCGGTGCTTGAATTTGTCAGGAGGCTTCTGCTGCTTTCGAACTGGCTTGT           | 1977 |
| Qy | 1936 | CACCTATTTCTGCTGCGCTGATTAAGACCAAAATGT                              | 1973 |
| Db | 1978 | CACCTGGTCCAGTCAAGAAAGACTACCAAGATATGT                              | 2015 |

RESULT 4  
US-08-459-569-8  
Sequence 8, Application US/08459569  
Patent No. 5543501  
GENERAL INFORMATION:  
APPLICANT: Slightom, Jerry L.  
APPLICANT: Tepfer, David A.  
TITLE OF INVENTION: R1 T-DNA Promoters  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GRAY, CARV, AMES & FRYE  
STREET: 401 B Street, Suite 1700  
CITY: San Diego  
STATE: California

```

1 COUNTRY: USA
2 ZIP: 92101-4297
3
4 COMPUTER READABLE FORM:
5
6 MEDIUM TYPE: floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: Patent In Release #1.0, Version #1.25
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/459,569
13 FILING DATE: 02-JUN-1995
14 CLASSIFICATION: 530
15
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 08/008,216
18 FILING DATE: 25-JAN-1993
19 APPLICATION NUMBER: US 06/725,368
20 FILING DATE: 22-APR-1985
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Barnhorst, Marnie W.
23 REGISTRATION NUMBER: 36,740
24 REFERENCE/DOCKET NUMBER: P1020U51
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (619) 699-2700
27 TELEFAX: (619) 236-1048
28 INFORMATION FOR SEQ ID NO: 8:
29
30 SEQUENCE CHARACTERISTICS:
31
32 LENGTH: 2279 base pairs
33 TYPE: nucleic acid
34 STRANDEDNESS: double
35 TOPOLOGY: linear
36 MOLECULE TYPE: DNA (genomic)
37
38 ORIGINAL SOURCE:
39 ORGANISM: Agrobacterium rhizogenes
40 STRAIN: Strain A4
41 IMMEDIATE SOURCE:
42 LIBRARY: Convolvulus arvensis plant cells
43 CLONE: Clone 7
44
45 FEATURE:
46
47 NAME/KEY: misc_feature
48 LOCATION: 1..2279
49
50 OTHER INFORMATION: /label= ORF8
51
52 Patent No. 5543501 /note= "Sequence ORF8 corresponds to bases 6609
53
54 OTHER INFORMATION: through 8888 of Seq. ID No. 5543501 19. "
55
56 US-08-459-569-8

```

|    | Query Match           | 6.0% ; Score 136.4 ; DB 1 ;                                    | Length 2279 ;        |
|----|-----------------------|--|----------------------|
|    | Best Local Similarity | 45.9% ; Pred. No. 8.7e-37 ;                                    |                      |
|    | Matches 623 ;         | Conservative 0 ; Mismatches 721 ;                              | Indels 14 ; Gaps 4 ; |
| QY | 623                   | TCGACCTTGCTCTACGACTGACGACCGCTTTTGTGACCAATGTTCCGATAGTGGACGATCG  | 682                  |
| Db | 665                   | TTGACCTGCGGGTACGACTATACAGGTTTTTTCCTCCAGAGTGGATGCCATGATGGTGG    | 724                  |
| QY | 683                   | GCTTCTTTCCGGAGAGATGTTCTTAAGCCGAAAGTGGCGGTCAATTGGCGCTGGCATTTCCG | 742                  |
| Db | 725                   | GGCATTTTCCGGGAAGCGCACACTACTCTAGTGTTCATCGTAGGAGCGCGGCGTCTCG     | 784                  |
| QY | 743                   | GACTCGGGTGGGCAACGAACCTGCTTCATCGTGGGGGTAGACAGATTTACATATATAGAG   | 802                  |
| Db | 785                   | GCTTGTGTGCTGCACAGAACTACTTGGCGGTGGCGTCAAGGAATACCTCTTTTCGATA     | 844                  |
| QY | 803                   | CAAGTGAATCGTGTGAGGCA---GCTTTGGTACATGCTTTAGGAGCGCTCTTAATG       | 859                  |
| Db | 845                   | CCGTGATGAGATCCGTAGTGTGGGGCAATGCCGATGCCAAAGCGGACGCTCACCAAG      | 904                  |
| QY | 860                   | TCGTGGCGGAATGGGGCGATGCGATTTCTCTCGTGCATCTGCTGTGTTTTCTTCC        | 919                  |
| Db | 905                   | CTTTAGCGTGTTCGGGTGTCATGCTTTCTCCGCAACCAACTTTGGCTGTCACTATAC      | 964                  |
| QY | 920                   | TTCAGCGTTACGGCGCTGTCTTCATGAGGCGCTTCCAAATCCCGGACCAATCGACATT     | 979                  |
| Db | 965                   | TGATTAAGTTTGAATTCGTCACGACCTTGCTGTGTTCTTGTGGCCGCACACACACACAG    | 1024                 |

QY 980 ACTTGCTCTACCAAGGCGCTCAATACATGTGAAAAGCCGGGACGTCGACCGAAGCTGT 1039  
 DB 1025 CACTATATTTCCGGCAGAAAACGCTACCATGCGACGGGGGCAAGCTCCGCGGGGATAT 1084  
 QY 1040 TCCATGCGCTTTACACAGGTTTGGCGTCCGTTCTTGAAGACGCTTTTCAATGACGAGATA 1099  
 DB 1085 TTGAGCGGGTACATGTGCGATGAGAACACTACTCTACCAAGGTGTGAACGGAATGCA 1144  
 QY 1100 TTGTGTGCTTGGCGCTGTGCTATTTACTCAGGCGCTTGAATCAGACACATTAGTGGG 1159  
 DB 1145 GGAGACTGATGCTCCGATGATATCTCTTTCATGTTGAAGACCGCTGCTGATGAG 1204  
 QY 1160 CTCATGATCTCTGCGCAAAATTTGGCTGAAACCGTTTGGAGAGAGTCTTCTTTCAGGA 1219  
 DB 1205 CCTCAGAAAGCAGCGGACGCTTGGCTCCGAGAGTTCCGAAATTCATTCTTCATGCGCTTT 1264  
 QY 1220 TAAAGAGATCTTTCTTGGGACACATCTCTGTTGT---AAACATGAGTTTCTC 1275  
 DB 1265 TGGTCGAGATCTTCACTGTGTGTAATTCAGTCTCTGTGGCAAGGCGTGGCAAAACACCCC 1324  
 QY 1276 ATGATTGGGACCTATTCAAGCTATGGAATAGGATCTGCGGGGTTTGGTCCAGTTT 1335  
 DB 1325 ATGATTTTCAGGCTTTGGGATCTGAGTTGGGATACGGCCGATTTCTGCTC-TATTAC 1383  
 QY 1336 GAAAGCGGGTTTATATGATCTCTCCGCTTGTATCAACGGAATGAGAAAATCAGCGG 1395  
 DB 1384 AACGTGTGTTTCAACGATCTGAGATGATATCAATGGTATCAGAGAGACAGCAT 1443  
 QY 1396 ATGTGCGCTGAAGAAATCTCAGAACTTCCACGTGGAATGCGATCTGAAAGTTAACGCT 1455  
 DB 1444 CTTTCTATTTGTTGGGGTTCACTTTTGCAGGCTCTGATGCGCATGTAATATTCAGAAA 1503  
 QY 1456 GTGTCTGTGAGCGACGACATATGCAATGTTCAAGTCAGGCGGATTCAGAAAAGACA 1515  
 DB 1504 AGCCATGCGGAACACGACACTGTTTGTGATCCCGTGTGTGGAATAGCAGAGAGGCGGG 1563  
 QY 1516 AAAATAAGATAAGCTTAAAGCGGGAATCTGAACTTTATGATAAGTGTGTGTACA 1575  
 DB 1564 AGATTGAAGGTATGTTGAAACACGCGCATTCGCGTGTTTTGAACGAGTCATCATGCGC 1623  
 QY 1576 TCTGAGCTGCAAAATCTCAACTCAGGATTTGCTGACATGCGATACCAATATTTTCAG 1635  
 DB 1624 GGCAGTGTGAGGCGCGCTACAGTTGATTAACAGACTGCGCGGATGAGACTTCTTCAGC 1683  
 QY 1636 GCACCAAGTGAACCAAGGCGTTGATTAACAGCCATATGACAGATGCTCAAACTTCTCCTG 1695  
 DB 1684 TACAATATGCAACCCGCGCTGCGAACTGCTGCGGCTGCAATTCAGACACTTTCATG 1743  
 QY 1696 ATGACTGAACGAAAATTTGTGTTAGACCAATATCTCCGCTTGTGTCTCATGACGGG 1755  
 DB 1744 GTCACGAAGCAAAAGTTTGGTTAACTCCGCGATCCGACAGTATGATGACCGATGGG 1803  
 QY 1756 ATGCGAAAAGCAATGATTTGCTGAGCTATGATGCGAGATTCGATGTTAAAGTCTA 1815  
 DB 1804 CTGTGCTGTAGCTGTGTGATGATGACATGAAATG-----CCAGCTGAGAGGGGCTT 1857  
 QY 1816 GTGCTCATCATGTTATCATGAGAGAGCACTCCCAAGCTGTTGGCGGCTCCCGACAA 1875  
 DB 1858 GTGCTTTTACATGATGTTTGAATGATATCTATCCCGCGCATCAGACATCATGACAG 1917  
 QY 1876 AAAGACGATTAATGTCTGTGCGGAGACGAATTTGAGATCTTTCCCGGCGTTGCGCAG 1935  
 DB 1918 AAGGACGCGTGTGGAATGATGTCAGGAGACTGTGCTGCTCTTCTGTAACGCGCTTGT 1977  
 QY 1936 CACTATTTTCTGCTGCGCTGATTTAGACCAAAATGT 1973  
 DB 1978 CACTGTGTCCAGTCAACGAAGCATGATATGT 2015

RESULT 5  
 US-08-458-831-8  
 ; Sequence 8, Application US/08458831  
 ; Patent No. 5824866

; GENERAL INFORMATION:  
 ; APPLICANT: Slightom, Jerry L.  
 ; APPLICANT: Tepfer, David A.  
 ; TITLE OF INVENTION: R1 T-DNA Promoters  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GRAY, CARV, AMES & FRYE  
 ; STREET: 401 B Street, Suite 1700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 92101-4297  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/458,831  
 ; FILING DATE: 02-JUN-1995  
 ; CLASSIFICATION: 800  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/008,216  
 ; FILING DATE: 25-JAN-1993  
 ; APPLICATION NUMBER: US 06/725,368  
 ; FILING DATE: 22-APR-1985  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Barnhorst, Marlene W.  
 ; REGISTRATION NUMBER: 36,740  
 ; REFERENCE/DOCKET NUMBER: P1020US1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (619) 699-2700  
 ; TELEFAX: (619) 236-1048  
 ; INFORMATION FOR SEQ. ID NO.: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2279 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; ORGANISM: Agrobacterium rhizogenes  
 ; STRAIN: Strain A4  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: Convulvulus arvensis plant cells  
 ; CLONE: Clone 7  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: 1..2279  
 ; OTHER INFORMATION: /label=ORF8  
 ; OTHER INFORMATION: /note="Sequence ORF8 corresponds to bases 6609  
 ; Patent No. 5824866  
 ; OTHER INFORMATION: through 8888 of Seq. ID No. 5824866 19. "  
 ; US-08-458-831-8

Query Match 6.0%; Score 136.4; DB 1; Length 2279;  
 Best Local Similarity 45.9%; Pred. No. 8,7e-37;  
 Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;

QY 623 TGACTTGTCTACGACGTCAGACCGTTTGTGACCAATGTTCCGATGATGACGAGATG 682  
 DB 665 TTGACCTGCGGACGACTATACAGTTTCTCCAGAGTGCATGCCATGATGTGTGG 724  
 QY 683 GCTTCTTCCGAGAGATGTTCTTAAGCCGAAAGTGGCGGCTAFTGGCGCTGGCATTTCCG 742  
 DB 725 GGCATTATCCGAAGGCGACACCTACCTAGTGTGCAATCGTAGAGGCGGGCTGTG 784  
 QY 743 GACTCGTGTGCAACGAAGTCTTCATGCTGGGTAGACGATGTTACATATATGAAG 802  
 DB 785 GCCTTGTGTGCTCCACAGAACTACTTGGCGCTGGGTGCAAGGAATCACTCTTTTGATA 844  
 QY 803 CAAGTGATCGTGTGAGGCA--GCTTGTGATCATGCTTTACAGGACGCTCTAGTG 859

Db 845 CCGTGTATGATCCGATGTTTGGGCGATCGCGATGCCAAACGGCGACGCTACACAG 904  
Qy 860 TCGTGCCGAATAGGGGCGATGATTTCTCTCGTGCATTCGTGTTTCTTCC 919  
Db 905 CTTTGACGTGTTCCGTGTATGCTTTCTCCGCAACCACTTTGCTGTCACTATC 964  
Qy 920 TCGAGCGTTACGGCTGTCTTTCATGATGAGCCGTTCCCAATCCCGGACAGTGCACATT 979  
Db 965 TGGATTAAGTTTAAGATTCGTCACAGCCTTGTTTCTTGTGCGGGAAGACACACAG 1024  
Qy 980 ACTTGCTACCAAGGGCTCAATACATGTGAAAGCCGGGACGCTCCAGCAAGCTGT 1039  
Db 1025 CACTATATTTTCCCGCAAGACGCTACGCAAGCGGGGCAAGCTCCCGGGATAT 1084  
Qy 1040 TCCATGGCGTTTCAACGGTTGCGTGTCTTGAAGACGGTTTTCATGACCGAGTA 1099  
Db 1085 TTACGGGGGACATGTGATGAGAACACTACTCTACCAAGGTGTGAACGGAATGGCA 1144  
Qy 1100 TTGTGTGGCTTCGCTGTCTTATTAAGCGCTTGAATCAGACACATTAGGTGG 1159  
Db 1145 GGAGACTGATGCTCCGATGATATCTTTCAATGTTGAAGACGCTCGTGTGATGAAG 1204  
Qy 1160 CTCATACATCTCTGCAAAATTTGCTGAACCGTTTCCGGAGAGCTCTTCTTCAAGGA 1219  
Db 1205 CCTCAGAGCAGCGCAGCTTGGCTCGAAGATTCCGAAATTCACCTTTCATGCCGTTT 1264  
Qy 1220 TTAGAGGATCTTCTGCGCACACATCTCTCGTGTGT---AAATGAGATTTTCTTC 1275  
Db 1265 TGGTCAGATCTTCAGCTGTGTAAATTCAGTCTTGTGCAAGGATGCAAAACACCCC 1324  
Qy 1276 ATGATTTGGACCTATTCAGCTATGAGGATAGGATCTGCGGGGTTTGGTCCAGTTT 1335  
Db 1325 ATGATTTTCAGGGCTTTCGGATATCTAGGTGGGATACGGCCGATTTCTGTC-TATTAC 1383  
Qy 1336 GAAAGCGGTTTATGATGATCTCCCTGCTGTGATCAACGATATGAAGAAATCAGCGG 1395  
Db 1384 AACGTGTTGTTTCAACGATCTCGACTGATTAATCAATGGTACAGAGGAGCAGCAT 1443  
Qy 1396 ATGTGCTGGAAGGATCTCAGAACTTCCAGCTCGATGCGATCTGAAGGTGTAACGGT 1455  
Db 1444 CTTTCTATTGTTGGGGTTCAACTTTTGCAGGCTGTATGCGCATTAATATTCGAAAA 1503  
Qy 1456 GTGTCTGTGAGCCAGCGCATATGCAATGTTCAAGTCAGGGCGATTCAGAAAGAAAA 1515  
Db 1504 AGCATGCGAAAGCAAGACTCTTTTGTATCCGTGCTGGAATACGCAAGGAGGCGGG 1563  
Qy 1516 AAAATTAAGATAAGCTTAAGAGCGGATATCTGAACCTTATGATGAAGTGTGTGACA 1575  
Db 1564 AGATTAAGGTATGTTGAACACGGTCAATTCGCTGTGTTTGAACAGGTCATCATTTGGC 1623  
Qy 1576 TCTGACTGGCAATATCCAACTCAGGCAATGCTGTGACATGCAATATTTTTCAG 1635  
Db 1624 GGCAGTGTGAGGCCCTCAAGTTGATTAACAGCACTGCGGGGATCAGACTTCTTCAGC 1683  
Qy 1636 GCACGATGAACCAAGCGGTTGATTAACAGCATATGACAGATGTCAAACCTTCTCTG 1695  
Db 1684 TACAAATATGAATCCGCGCTCGAAACTGCTGCGCTGTCAATTCAGACTTTCATG 1743  
Qy 1696 ATGACTGAAGCAAAATTTGTTAGACCAATCTCCGCTTGTGTCTCATGACGGG 1755  
Db 1744 GTACAGAGCAAAAGTTTGGGTTAATCCCGGATCCAGCAATGATGAACGATGGG 1803  
Qy 1756 ATTCGCAAAAGCAATGATGCTGAGTATGATGCGAGATCGAATGTAAGTCTA 1815  
Db 1804 CTTGTCGTGAGCTGTGTTGATTAACATGATG-----CCTGAGAGGGGCTT 1857  
Qy 1816 GTGCTCATGATATACATGAGGAGAGCACTCCCAAGAGTGTGCGGCGCCGACAA 1875  
Db 1858 GTGCTTTTTCACATGCTTGTGATGATCTATTCGCGGCGATCGAGCATATGCAAG 1917  
Qy 1876 AAGAGCGATATATGTGTGCTGCGGAGCAATTTGAGATCTTTCCCGCGTTTGCAG 1935  
Db 1918 AAGGAGCGGTGTGAAATTTGTCAGGAGCTTGTGCTGCTTCTGTAAGTGGCTTGT 1977

Qy 1936 CACCTATTTCTGCTGCGCTGATTAAGACCAAAATGT 1973  
Db 1978 CACTGTGCTCCAGTCAAGCAAGCTACGAACTATATGT 2015  
  
RESULT 6  
US-08-008-216-19  
Sequence 19, Application US/08008216  
Patent No. 5366887  
GENERAL INFORMATION:  
APPLICANT: Slightom, Jerry L.  
APPLICANT: Tepfer, David A.  
TITLE OF INVENTION: R1 T-DNA Promoters  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GRAY, CARY, AMES & FRYE  
STREET: 401 B Street, Suite 1700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-4297  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/008,216  
FILING DATE: 25-JAN-1993  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 06/725,368  
FILING DATE: 22-APR-1985  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhore, Marlie W.  
REGISTRATION NUMBER: 36,740  
REFERENCE/DOCKET NUMBER: P1020US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 699-2700  
TELEFAX: (619) 236-1048  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21126 base pairs  
TYPE: nucleic acid  
TOPOLOGY: linear  
STRANDEDNESS: single  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Agrobacterium rhizogenes  
STRAIN: STRAIN A4  
LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS  
CLONE: CLONE 7  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (937..2262)  
OTHER INFORMATION: /label= ORF1SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (2649..3458)  
OTHER INFORMATION: /label= ORF2SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3726..4799  
OTHER INFORMATION: /label= ORF3SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (4041..4400)  
OTHER INFORMATION: /label= ORF4SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature

```
/ LOCATION: complement (4607..4918)
/ OTHER INFORMATION: /label= ORF5SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5143..6216
/ OTHER INFORMATION: /label= ORF6SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (5071..5643)
/ OTHER INFORMATION: /label= ORF7SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6609..8888
/ OTHER INFORMATION: /label= ORF8SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (6576..6830)
/ OTHER INFORMATION: /label= ORF9SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9748..10044
/ OTHER INFORMATION: /label= ORF10SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (10509..11282)
/ OTHER INFORMATION: /label= ORF11SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 12466..13002
/ OTHER INFORMATION: /label= ORF12SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 13723..14319
/ OTHER INFORMATION: /label= ORF13SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 15659..16210
/ OTHER INFORMATION: /label= ORF14SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (16517..17545)
/ OTHER INFORMATION: /label= ORF15SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (17737..18189)
/ OTHER INFORMATION: /label= ORF16SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (18177..18743)
/ OTHER INFORMATION: /label= ORF17SUBSEQUENCE
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: complement (19031..19390)
/ OTHER INFORMATION: /label= ORF18SUBSEQUENCE
/ US-08-008-216-19

Query Match 6.0%; Score 136.4; DB 1; Length 21126;
Best Local Similarity 45.9%; Pred. No. 4.8e-36;
Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;

QY TCGACTGCTCTACGACTGCGAGACCGTTTTCATGATGTCGATAGTGAAGCGATCG 682
DB TTTGACCTGCGGTACGATCTATCAGGTTTCTCCAGAGTGTGATGCCCATGATGTTGG 7332
QY 683 GCTTCTTTCGCGAGATGTTCTTAAGCCGAAAGTGGCGGTCAATTGGCGCTGGCATTTCCG 742
DB 7333 GGCATTATCCGGAAGGCGCACACTACTAGTGTTCATGCTGTGAGAGCGGCGCTGTG 7392
QY 743 GACTCGGTGGGCAAGAACTGTTTATGCTGTGGGGTGAAGAGATGTTACATATATAG 802
DB 7393 GCCTTGTGTGCGCAGAACTACTGTGGCGTGGCGTCAAGGAATACCTTTTCGATA 7452
QY 803 CAAGTATCGTGTGAGGCA--GCTTGTGCATGCTTTTCAAGGAGCGTCTAGTG 859
```

```
DB 7453 CCGTATAGATCCGTAATTTTGGGGCATCCGCCATGCCAAACGGCGACGCTCACCAAG 7512
QY 860 TCGTGGCCGAATATGGGGCGGATGCCATTTCTTCCTGCTGATCTTCCTTTTCTTC 919
DB 7513 CTTGACGTGTGCTGTGATGCTTTCTTCGCGCAACCAACTTGTGCTATCTATCT 7572
QY 920 TCGAGGCTTACGAGCCGTCTTCGATGATGAGCCGTTCCCAATCCGCGACAGTGCACATT 979
DB 7573 TGGATTAAGTTTAAATTCCTGCTCAGCCTTCGTTTCTGTGTGCGGCAACGACACAG 7632
QY 980 ACTTGTCTACCAAGGCGTCAATACATGTGGAACCCGGGACGCTGCCACGACCTGT 1039
DB 7633 CACTATATTTCCGCCAGAAACGCTACGATGCGACGCGGGCAAGCTCCGCGGGATAT 7692
QY 1040 TCCATGCGCTTTTAAACAGGTTGCGCGTCTTCTGAAGAACGTTTTCATGACGAGATA 1099
DB 7693 TTCAGCGGTACATGTGGATGGAACACTACTTACCAAGGTTGAAACGAAATGGCA 7752
QY 1100 TTGTGTGGCTTCGCTGCTGCTTACTCAGGCTTGAATCAGGACATTAAGGTGG 1159
DB 7753 GGAAGACTGATGGCTCCGATGATATCTCTTCAATGTTAAAGCGCTCGCTGATGAAG 7812
QY 1160 CTCATGACTCTGCGCAAAATTTGGCTGAACCGTTTCGGAAGGAGTCTTCTTCAAGGA 1219
DB 7813 CTTCAAGACACGCGACGCTTTGGCTCCGAGAGTTCGAAATTCATCTTCCAGCGTTT 7872
QY 1220 TAGAGAGATCTTCTTCGCGACACATCTCTGCTGTGTTT---AAATGAGTTCCTC 1275
DB 7873 TGGTCAGATCTTCACTGCTGTGTAATTCAGTCTGTGTGCAAGGATGCAACCAACCCC 7932
QY 1276 ATGATTGGACCTTATTCAGCTAATGGAATAGATCTGCGCGGTTTGGTCCAGTTT 1335
DB 7933 ATGATTTTCAAGGCTTTCCGGATATCTAGGTGTGGGATACGGCCAGATTGCTC-TATTAC 7991
QY 1336 GAAAGGCGTTTATAGATCTCTCGCTGTGTATCAACGATATGAAGAAATACGCGG 1395
DB 7992 AAGGTGTGTTTCAACGATCTGGAATGATGATTAATGCTTACAGAGGACACGACAT 8051
QY 1396 ATGTGCCCTGAAGGAATCTCAAGACTTCCAGCTGGATGCGATCTGTAAGCTTAACGT 1455
DB 8052 CTTTCTATTGTGGGGTCAACTTTTGACAGCTCTGATGCGCAATTAATATTCAGAAA 8111
QY 1456 GTGTCTGTAGCGCGCATATGSCATGTTCAAGTCAGGCGCATTCAGAAAGAAAGACA 1515
DB 8112 AGCATTGCGAAAGCAAGACTCTTTTGTATCCGTCGTGGAATTAACCAAGAGGCGGG 8171
QY 1516 AAAATTAAGATTAAGGCTTAAGACGGGATATCTGAACCTTTATGATTAAGTGTGTACA 1575
DB 8172 AGATTGAAGTATGTTGAACAACGGTCAATTCGCTGTTTGAACCAAGTCATCATTTGCG 8231
QY 1576 TCTGACTGCGAAATATCCAACTCAGGCAATGCTTCATGATGCAATATATTTTCAG 1635
DB 8232 GGCAGTGTGAGGCGCTCACTACATGTAATACAGCTGCGGGATAGATATCTTCTTACG 8291
QY 1636 GCACCAAGTAACCAAGCGGTGATTAACAGCCATATGACAGATCGTCAAAACTCTTCTG 1695
DB 8292 TACAAATATGAAACCCGCGTCCGAAACTGCTGCGCTGTCAATTCAGACCTTTCATG 8351
QY 1696 ATGACTGAACGAATAATCTGTGTTAGAACATATCTCCGCTTGTGCTCATGACGGG 1755
DB 8352 GTCACGAGCAAAAGTTTGGGTTAACTCCGGCATCCAGCAATGATATGACGATGGG 8411
QY 1756 ATGCAAAAGCAATGATATGCTGAGCTATGATGATGCGAGATCCGAATGTTAAAGCTTA 1815
DB 8412 CTGTGCTGAGGCTGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 8465
QY 1816 GTGCTATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1875
DB 8466 GTGCTTTTACATATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATG 8525
QY 1876 AAGAGCGATTAATGCTGTGCTGCGGAGCGCAATTTGAGATCTTTCGCGGGTTCGCGAG 1935
```

Db 8526 AAGGACGCGTGCCTGGATTGGTCAAGGAGAGCTGCTGCTTCTGAACTGGCTTGT 8586

Qy 1936 CACCTATTTCCTGCCTGGCTGATTACGACCAAAATGT 1973

Db 8586 CACCTGGTCCAGTCAAGCAAGACTGACGACGATATGT 8623

## RESULT 7

```

US-08-459-569-19
: Sequence 19, Application US/08459569
: Patent No. 5543501
: GENERAL INFORMATION:
: APPLICANT: Slightom, Jerry L.
: APPLICANT: Tepfer, David A.
: TITLE OF INVENTION: R1 T-DNA Promoters
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GRAY, CARY, AMES & FRIE
: STREET: 401 B Street, Suite 1700
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92101-4297
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/459,569
: FILING DATE: 02-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/008,216
: FILING DATE: 25-JAN-1993
: APPLICATION NUMBER: US 06/725,368
: FILING DATE: 22-APR-1985
: ATTORNEY/AGENT INFORMATION:
: NAME: Barnhorst, Marlene W.
: REGISTRATION NUMBER: 36,740
: REFERENCE/DOCKET NUMBER: P1020U51
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 699-2700
: TELEFAX: (619) 236-1048
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 21126 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ORIGINAL SOURCE:
: ORGANISM: Agrobacterium rhizogenes
: STRAIN: STRAIN A4
: IMMEDIATE SOURCE:
: LIBRARY: CONVOLUTULUS ARVENSIS PLANT CELLS
: CLONE: CLONE 7
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: complement (937..2262)
: OTHER INFORMATION: /label=ORF1SUBSEQUENCE
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: complement (2649..3458)
: OTHER INFORMATION: /label=ORF2SUBSEQUENCE
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 3726..4799
: OTHER INFORMATION: /label=ORF3SUBSEQUENCE
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: complement (4041..4400)

```

```

1  OTHER INFORMATION: /label= ORF4SUBSEQUENCE
2  FEATURE:
3  NAME/KEY: misc_feature
4  LOCATION: complement (4607..4918)
5  OTHER INFORMATION: /label= ORF5SUBSEQUENCE
6  FEATURE:
7  NAME/KEY: misc_feature
8  LOCATION: 5143..6216
9  OTHER INFORMATION: /label= ORF6SUBSEQUENCE
10 FEATURE:
11 NAME/KEY: misc_feature
12 LOCATION: complement (5071..5643)
13 OTHER INFORMATION: /label= ORF7SUBSEQUENCE
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: 6609..8888
17 OTHER INFORMATION: /label= ORF8SUBSEQUENCE
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: complement (6576..6830)
21 OTHER INFORMATION: /label= ORF9SUBSEQUENCE
22 FEATURE:
23 NAME/KEY: misc_feature
24 LOCATION: 9748..10044
25 OTHER INFORMATION: /label= ORF10SUBSEQUENCE
26 FEATURE:
27 NAME/KEY: misc_feature
28 LOCATION: complement (10509..11282)
29 OTHER INFORMATION: /label= ORF11SUBSEQUENCE
30 FEATURE:
31 NAME/KEY: misc_feature
32 LOCATION: 12466..13002
33 OTHER INFORMATION: /label= ORF12SUBSEQUENCE
34 FEATURE:
35 NAME/KEY: misc_feature
36 LOCATION: 13723..14319
37 OTHER INFORMATION: /label= ORF13SUBSEQUENCE
38 FEATURE:
39 NAME/KEY: misc_feature
40 LOCATION: 15659..16210
41 OTHER INFORMATION: /label= ORF14SUBSEQUENCE
42 FEATURE:
43 NAME/KEY: misc_feature
44 LOCATION: complement (16517..17545)
45 OTHER INFORMATION: /label= ORF15SUBSEQUENCE
46 FEATURE:
47 NAME/KEY: misc_feature
48 LOCATION: complement (17737..18189)
49 OTHER INFORMATION: /label= ORF16SUBSEQUENCE
50 FEATURE:
51 NAME/KEY: misc_feature
52 LOCATION: complement (18177..18743)
53 OTHER INFORMATION: /label= ORF17SUBSEQUENCE
54 FEATURE:
55 NAME/KEY: misc_feature
56 LOCATION: complement (19031..19390)
57 OTHER INFORMATION: /label= ORF18SUBSEQUENCE
58 JS-08-459-569-19

```

|                           |        |                    |            |               |
|---------------------------|--------|--------------------|------------|---------------|
| Query Match               | 6.0%;  | Score 136.4;       | DB 1;      | Length 21126; |
| Best Local Similarity     | 45.9%; | Pred. No. 4.8e-36; |            |               |
| Matches 623; Conservative | 0;     | Mismatches 721;    | Indels 14; | Gaps 4;       |

|    |      |   |      |
|----|------|---|------|
| Qy | 623  | TCGACCTGGCTCTACGACTGCGAGACCGTTTTTTTGACCAATGTTCCGATATAGTGAAGGATTCG | 682  |
| Db | 7273 | TTGACCTGCGGTACGACTTATCAGGTTTTCTCCAGAGTGGATGCCCATGATGTGTGG         | 7332 |
| Qy | 683  | GCTTCCTTCCGAGAGATGTTCTTAAGCCGAAAGGCGGTATCGGCGCTGGCATTTCCG         | 742  |
| Db | 7333 | GGCATTTATCCGGAAAGGCGCACACTACTAGTGTGGCATGTGTAGAGAGCGGCGCTGTCTG     | 7392 |
| Qy | 743  | GACTCGTGTGGCGCAACGAACTGTTTCATGCTGGGGTAGACGATGTTACATATATGAG        | 802  |

Db 7393 GCCTGTGCTGCACAGAACTACTTGGCGCTGGCGTCAAGAAATCACTCTTTGGATA 7452  
Qy 803 CAGGTATGCTGTGGAGCAA---GCTTGTGCATGCTTTGAGGAGCGTCTAGTG 859  
Db 7453 CCCTTATGAGATCCGATGTTTGGGGCATCGCGATGCCAAACGGGAGGCTCACAGG 7512  
Qy 860 TCGTGCCGGAATGGGGCGCATGCTTCTCTGCTGCATTTCTGCTTTTCTTCC 919  
Db 7513 CTTGACGTCGTTGGGTGATGCTTTCTCCGCAACCACTTTGCTGTCACTATC 7572  
Qy 920 TCGAGCGTTACGGGCTGTCTTCAATGAGCGCTTCCCAATCCCGCACATGACATT 979  
Db 7573 TGTATAGTTTAAATATCCCTCCAGCCTTGTCTTGTGCGGGAACAGCACACAG 7632  
Qy 980 ACTTGTCTACCAAGGCGTCAATATGAGAAAGCGGGGAGCTGCCACGAGCTGT 1039  
Db 7633 CACTATATTTCCCGCAGAAACGCTACCATGACCGGGGCAAGCTCCCGGGATAT 7692  
Qy 1040 TCCATGCGCTTTTACACGCTTGGCGCTGCTTCTTGAAGACGGTTTTCATGACGATA 1099  
Db 7693 TTCAAGGGGTAATGTCGGATGGAAGACACTACTACCAAGGTGTGAACGGAATGCA 7752  
Qy 1100 TTGCTGTGCTTGGCTGCTGCTATTACTGAGCGCTTGAATACGACATTAAGTGG 1159  
Db 7753 GGAGACTGATGGCTCCGATGATATCTTTCAATGTTGAAGACCGTCTGATGTAAG 7812  
Qy 1160 CTGATGCTCTGGCAAAATTTGGCTGACCGTTCCGGAGAGGAGTCTTCTTACGGA 1219  
Db 7813 CCTCAGAGACGCGACGCTTGGCTCCGAGATTCGGAATAATCACTTTCACGCGTTT 7872  
Qy 1220 TAGAGAGATCTTTTGGGACACATCTCTGGTGT---AAACATGAGTTTCTC 1275  
Db 7873 TGGTCAAGATCTTCACTGTGTGTAATTCAGTCTGTGACAGCGATGCAACACCCC 7932  
Qy 1276 ATGATGGGACCTATTACAGCTAATGGAATAGAGCTGCGGGGTTTGTCCAGTTT 1335  
Db 7933 ATGATTTGAGGCTTTCGGGATCTAGTGTGGATACGGCCGAGTTTCTTCC-1ATATAC 7991  
Qy 1336 GAAAGCGGTTTATGATCTCCGCTGTGATCAACAGGATATGAAGAAATACGCGG 1395  
Db 7992 AAGGTTGTTTTCACGATCTGAGCTGATTAATCATGCTACGAGGAGACCGCAT 8051  
Qy 1396 ATGTGCGCTGAAAGATCTCAGAACTTCCAGCTGAGTCCCATCTGAAGTGTAAAGT 1455  
Db 8052 CTTTCAATGTGTGGGTTCACTTTTGCAGGCTGTGATGGGCAATGAAATATTCAGAAA 8111  
Qy 1456 GTGTCTGTGACGAGCGCATATGCCATGTTCAAGTCAGGCGATTCAAGAGAAAGACA 1515  
Db 8112 AGCCATGCGAAAGACGACTCTGTTTGAATCCGCTGCGAAATAGCCAAAGAGGCGGG 8171  
Qy 1516 AAAATTAAGTAAAGCTTAAGCGGGGATATCTGAACTTTATGATTAAGTGTGTGACA 1575  
Db 8172 AGATTAAGGTATGCTTGAACACGGGCAATTCGGGTGTTTGAACGAGTCAATGCGC 8231  
Qy 1576 TCTGACTCCCAATATCAACTAGGCAATGCTGACATGCGATACCAATATTTTTCAG 1635  
Db 8232 GGCAGTCTAGGCGCGCTACAGTTGAATACAGACTGCGCGGATGAGACTTCTTACG 8291  
Qy 1636 GCACCACTGAACCAAGGCTTGAATACAGCCATATGAACAGATCTCAAACTTCTTG 1695  
Db 8292 TACAAATTCGAACCCGCGCTCGAAACTCGTCTGCGCTCAATCAGACCTTCTATG 8351  
Qy 1696 ATGACTGAAGAAATTTCTGTTAGACCATATCTCCCGTCTTGTCTTCATGAGAGGG 1755  
Db 8352 GTACAGGAAGAAAGTTTGGTTAACTCCGGCAATCCAGAGTGAATGAGACGAGTGG 8411  
Qy 1756 ATGCAAAACAGATGTAATGCTGCTATGAGTCGACAGATCCGAATGTAAGGCTTA 1815  
Db 8412 CTGTGCTGAGCTGTGCTTGAATGATGAAATG---CCAGTGAAGAGGCGCTT 8465  
Qy 1816 GTGCTCATCATTAATACATGGAAGACGACTCCCAAGCTGTGTGGCGGTCCCGACAAA 1875  
Db 8466 GTGCTTTTACATATGCTTTGATGATCTATCTATCCGGCGGATGAGCATATGACAAG 8525

Qy 1876 AAAGAGCATTAATGCTGCTGGGGAGCAATTTGAGATCTTTCCCGCGTTGCCAG 1935  
Db 8526 AAGAGCGGTGCTTGAATTTGTCAGGAGCGTTCGCTGCTCTTAACATGCGCTGT 8585  
Qy 1936 CACTATTTCTGCTGCGCTGATTAACGACAAATGT 1973  
Db 8586 CACTGTCCCACTCAACGAGACTACGAAAGATATGT 8623  
RESULT 8  
US-08-458-831-19  
; Sequence 19, Application US/08458831  
; Patent No. 5824866  
; GENERAL INFORMATION:  
; APPLICANT: Slightcom, Jerry L.  
; APPLICANT: Tepfer, David A.  
; TITLE OF INVENTION: R1 T-DNA Promoters  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: GRAY, CARY, AMES & FRYE  
; STREET: 401 B Street, Suite 1700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-4297  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/458, 831  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 800  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/008, 216  
; FILING DATE: 25-JAN-1993  
; APPLICATION NUMBER: US 06/725, 368  
; FILING DATE: 22-APR-1985  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barnhorst, Marilee W.  
; REGISTRATION NUMBER: 36,740  
; REFERENCE/DOCKET NUMBER: P1020US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 689-2700  
; TELEFAX: (619) 236-1048  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21126 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEtical: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Agrobacterium rhizogenes  
; STRAIN: STRAIN A4  
; IMMEDIATE SOURCE:  
; LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS  
; CLONE: CLONE 7  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: complement (937..2262)  
; OTHER INFORMATION: /label= ORF1SUBSEQUENCE  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: complement (2649..3458)  
; OTHER INFORMATION: /label= ORF2SUBSEQUENCE  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 3726..4799  
; OTHER INFORMATION: /label= ORF3SUBSEQUENCE

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (4041..4400)  
OTHER INFORMATION: /label= ORF4SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (4607..4918)  
OTHER INFORMATION: /label= ORF5SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 5143..6216  
OTHER INFORMATION: /label= ORF6SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (5071..5643)  
OTHER INFORMATION: /label= ORF7SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6609..8888  
OTHER INFORMATION: /label= ORF8SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (6576..6830)  
OTHER INFORMATION: /label= ORF9SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 9748..10044  
OTHER INFORMATION: /label= ORF10SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (10509..11282)  
OTHER INFORMATION: /label= ORF11SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 12466..13002  
OTHER INFORMATION: /label= ORF12SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 13723..14319  
OTHER INFORMATION: /label= ORF13SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 15659..16210  
OTHER INFORMATION: /label= ORF14SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (16517..17545)  
OTHER INFORMATION: /label= ORF15SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (17737..18189)  
OTHER INFORMATION: /label= ORF16SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (18177..18743)  
OTHER INFORMATION: /label= ORF17SUBSEQUENCE  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: complement (19031..19390)  
OTHER INFORMATION: /label= ORF18SUBSEQUENCE  
US-08-458-831-19

Query Match 6.0%; Score 136.4; DB 1; Length 21126;  
Best Local Similarity 45.9%; Pred. No. 4.8e-36;  
Matches 623; Conservative 0; Mismatches 721; Indels 14; Gaps 4;

QY 623 TCACCTTCTCTGCTGACGACCGCTTTTGTACCAATGTTCCGATGAGACGATG 682  
Db 7273 TTGACCTGCGGTACGACTATCAGTTTCTCTCAAGAGTGCATGCCCATGATGCTG 7332

QY 683 GCTTCTTTCGAGAGTGTCTTAAGCCGAAGTGGCGGTATGGCGCTGGCATTTCCG 742  
Db 7333 GGCATTATCCGGAAGGCGACCACTAGTGTGTCATCGTAGAGAGGGGCTGTCTG 7392

QY 743 GACTCGTGGGCAAGCACTGCTTCATGCTGGGGTAGACGATGTACATATATGAG 802  
Db 7393 GCCTTGTGCTGCACAGAACTACTTGGCGCTGGCGTCAAGAAATCACTTTTGCATA 7452

QY 803 CAAGTATCGTGTGGAGCA--GCTTGTACATGCTTTTCAGGACGCTCTAGT 859  
Db 7453 CCGTTATAGATCCGTAAGTTTGGGGCATCCGCCATGCCAAGCGGAGCGCTCACCAG 7512

QY 860 TCGTGCCGAATGGGGCGATGCAATTCCTCTGCTGATTCCTGTTTCTTC 919  
Db 7513 CTTTGACGTGCTGGGTGTCATGCTTTCGCCCAACCAACTTGTGCTGTCACTATC 7572

QY 920 TCAGAGGTTACGCGCTGTTCATGAGCGCTTCCCAATCCCGACAGTGACACTT 979  
Db 7573 TGGATAGTTTAATATTCGTCACGCTTGTTTCTTGTGCGGCAACGACACACAG 7632

QY 980 ACTTGTCTACCAAGCGTCCATATCATGTGAAACCGGCGACCTGCCACCAAGCTGT 1039  
Db 7633 CACTATATTTCCCGCAGAAACGCTACGCATGSCACCGGGGCAAGCTCCGCCGGATAT 7692

QY 1040 TCCATGCGTTTACAAACGTTGGCGTCTTCTTGAAGACGGTCTTCATGACGAGATA 1099  
Db 7693 TTGACGGGGTACATGTGCGATGGAGACACTACTTACCAAGGGTGTGAACGGAATGCA 7752

QY 1100 TTGTGTTGGCTTGGCTGTCTGCTATCTACAGGCTTGAATACGACATTAAGTGG 1159  
Db 7753 GGAGACTGATGGCTCGATGATATCTCTTCAATGTTGAAGACGCTGTGATGATAG 7812

QY 1160 CTCATGACTCTGCAAAATTTGGCTGACCGTTTCGGAGGAGAGTCTTCTTCAAGGA 1219  
Db 7813 CCTCAGAACAGCGACGCTTGTGCTCCGAGAGTTCGAAATTCATTTTCATCCGTTT 7872

QY 1220 TAGAGAGATCTTTCTGGGACACATCTCTGCTGTGCTTCTC 1275  
Db 7873 TGGTCAGATCTTCACTGTGATTTGAGTCTGTGGCAAGGATGCAACACCCC 7932

QY 1276 ATGATTGGGACCTATTCAGTAAATGGAATAGGATCTGGCGGGGTTTGTCCAGTTT 1335  
Db 7933 ATGATTTCAAGGCTTTCGGGATATCTGAGTTGGGATACGCCCAAGTTTGTCTC-7ATTTAC 7991

QY 1336 GAAAGCGGTTTATTGATCTCTCGCTTGTGATCAACGATATGAGAAATTCAGCGG 1395  
Db 7992 AACGTGTGTTTCAACGATCTCTGACGTGATATTAATGCTACAGAGACCAAGCAT 8051

QY 1396 ATGTGCTGAAAGATCTCAGAACTTCAAGCTGATCGCATCTGAAGTGTAAACGCT 1455  
Db 8052 CTTTCTATTGTTGGGTTCAACTTTTGACAGGCTGTATGCGCATGAAATATTCAGAA 8111

QY 1456 GTGTCTGTAGCCAGCGCATATGCCATGTTCAAGTCAAGGCCATTCAGAAAGAAAGACA 1515  
Db 8112 AGCCATGCAAAAGCAAGACTGTGTTTGAATCCGTCGCTGGAATACCAAGAGGGCGG 8171

QY 1516 AAAATTAAGATAGGCTTAAGACGGGATATCTGAACCTTATGATAAGTGGTGCACA 1575  
Db 8172 AGATTGAAGTATCTTGAACACGCTCATTTGCCGTTTGTGACCAAGCTCATATGGC 8231

QY 1576 TCTGACTGCAAAATATCCAACTCAGGACTTGCCTGACATGCAATCAATATTTTCAG 1635  
Db 8232 GGAGAGTGCAGAGCGGCTACAGTTGATTAACAAGCTGGCGGGATAGACTTCTTCAGC 8291

QY 1636 GCAACAGTAAACCAAGCGGTTGATTAACAGCCATATGACAGATGCTCAAACTTCTCTG 1695  
Db 8292 TACAATATGAACCCGCCGTGCAAACTGTGCGCTGTCAATTCAGACTCTTCATG 8351

QY 1696 ATGACTGAAGCAAAATTTCTGTTAGACCATATCCCTGCTTGTGCTCATAGGACGG 1755  
Db 8352 GTACAGAACAAAGTTTGTGGTTAACTCCGATCTCCAGCTGATATGACCCGATGG 8411

QY 1756 ATGCAAAAGCAAGTATATCTGACCTATGATGTCAGAGATCCGAATGAAGTCTTA 1815  
Db 8412 CTTGTCCGTAGACTGTGTTGACTTGAATGATCGAATG-----CAAGCTGAGAGGCGCTT 8465



SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,037  
FILING DATE: June 13, 1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/005,167  
FILING DATE: October 13, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 220/156  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1642 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 97...1446  
US-08-665-037-1

Query Match 1.6%; Score 36; DB 2; Length 1642;  
Best Local Similarity 62.0%; Fred. No. 0.15;  
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1814 TAGTCTCATCATGATTATACATGGAGAGAGACTCCCAAGCTGTGGCGGTCCCGACA 1873  
DB 728 TGTGGCCAGCATCAAGCAGCAGGAGAGAGAGCGCCGCTGCTGCTGTGAGACCCGACA 787  
QY 1874 AAAAAGCGGATTATGTCTGCTCGGAGACCA 1905  
DB 788 CAGATGAACACTTCAAGCGGCTTCGGGTCA 819

RESULT 13  
US-08-666-067-1  
Sequence 1, Application US/08666067  
Patent No. 5922842  
GENERAL INFORMATION:  
APPLICANT: Seedorf, Klaus  
APPLICANT: Ulirich, Axel  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT  
TITLE OF INVENTION: OF TKA-1 RELATED  
TITLE OF INVENTION: DISORDERS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/666,067  
FILING DATE: June 13, 1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/005,421  
FILING DATE: October 13, 1995  
ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 220/157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1642 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 97...1446  
US-08-666-067-1

Query Match 1.6%; Score 36; DB 2; Length 1642;  
Best Local Similarity 62.0%; Fred. No. 0.15;  
Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1814 TAGTCTCATCATGATTATACATGGAGAGAGACTCCCAAGCTGTGGCGGTCCCGACA 1873  
DB 728 TGTGGCCAGCATCAAGCAGCAGGAGAGAGAGCGCCGCTGCTGCTGTGAGACCCGACA 787  
QY 1874 AAAAAGCGGATTATGTCTGCTCGGAGACCA 1905  
DB 788 CAGATGAACACTTCAAGCGGCTTCGGGTCA 819

RESULT 14  
US-08-732-870-1  
Sequence 1, Application US/08732870  
Patent No. 5945523  
GENERAL INFORMATION:  
APPLICANT: Seedorf, Axel  
APPLICANT: Ulirich, Axel  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT  
TITLE OF INVENTION: OF TKA-1 RELATED  
TITLE OF INVENTION: DISORDERS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/732,870  
FILING DATE: October 15, 1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/666,037  
FILING DATE: June 13, 1996  
APPLICATION NUMBER: 08/666,067  
FILING DATE: June 13, 1996  
APPLICATION NUMBER: 60/005,167  
FILING DATE: October 13, 1995  
APPLICATION NUMBER: 60/005,421  
FILING DATE: October 13, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 222/247

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1642 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 97...1446  
US-08-732-870-1

Query Match 1.6%; Score 36; DB 2; Length 1642;

Best Local Similarity 62.0%; Pred. No. 0.15; Mismatches 35; Indels 0; Gaps 0;

Matches 57; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1814 TAGTGCTCATGATTATACATGGAGAGACGACTCCCAAGCTGTGGCGTCCCGACA 1873

DB 728 TGGTGGCCAGCATCAGGACGCGAGAGACAGGCTCCGCTGTGTGTGACCCGACA 787

QY 1874 AAAAGAGCGATTATGTCTGCTCGGAGCGCA 1905

DB 788 CAGATGACACTTCAGGCGCTTCGGGTACA 819

## RESULT 15

US-09-415-277C-4/C

Sequence 4, Application US/09415277C

Patent No. 6531308

GENERAL INFORMATION:

APPLICANT: Hershenberger, Charles

APPLICANT: Payson, Robert

TITLE OF INVENTION: Ketoreductase Gene and Protein from Yeast

FILE REFERENCE: X-11325A

CURRENT APPLICATION NUMBER: US/09/415,277C

PRIOR APPLICATION NUMBER: US 09/182,985

PRIOR FILING DATE: 1998-10-30

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4

LENGTH: 1032

TYPE: DNA

ORGANISM: S. cerevisiae

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(1032)

OTHER INFORMATION:

US-09-415-277C-4

Query Match 1.6%; Score 35.2; DB 4; Length 1032;

Best Local Similarity 53.7%; Pred. No. 0.2;

Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1498 ATTCAGAGGAAAGACAAAATAAGATTAGGCTTAAGCGGATATCTGAACCTTAT 1557

DB 624 ATCGAAAAGTTGAGGACCAAAAGAGAGGTTGATGTTGTTGATTGAT 565

QY 1558 GATAAGTGTGTCATCATCTGCGCAATATCCAACTCAGAGCATTTGCGTCAGATGC 1617

DB 564 GTGATCTTCATCTCTTTTGTGAACCTCCAGGAGCCTTTTCAGCAACTTCTTGATGC 505

QY 1618 GATACCAATATTTTC 1633

DB 504 AAGTAAGCATTTATC 489

Search completed: November 23, 2003, 21:26:57  
Job time: 147.658 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2003, 21:01:51 ; Search time 4831.41 Seconds  
(without alignments)  
11409.196 Million cell updates/sec

Title: US-09-434-837-10

Perfect score: 2268  
Sequence: 1 acgtcagcttcaccctctcct.....actaccgcagtagaattag 2268

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*  
1: em\_estda: \*  
2: em\_esthum: \*  
3: em\_estlin: \*  
4: em\_estm: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hlc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hlc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estcom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pln: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_man: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rtd: \*  
26: em\_gss\_phg: \*  
27: em\_gss\_vrl: \*  
28: gb\_gss1: \*  
29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 51.6  | 2.3         | 1201   | 13 | BX381961 BX381961 |
| 2          | 49.8  | 2.2         | 618    | 9  | AU239500 AU239500 |
| 3          | 44.8  | 2.0         | 564    | 12 | BJ123611 BJ123611 |
| 4          | 44.2  | 1.9         | 461    | 12 | BJ101485 BJ101485 |

|    |      |     |      |    |                   |
|----|------|-----|------|----|-------------------|
| 5  | 44.2 | 1.9 | 486  | 12 | BJ103567 BJ103567 |
| 6  | 44.2 | 1.9 | 508  | 9  | AU199996 AU199996 |
| 7  | 44.2 | 1.9 | 511  | 12 | BJ108542 BJ108542 |
| 8  | 44.2 | 1.9 | 513  | 9  | AU203802 AU203802 |
| 9  | 44.2 | 1.9 | 529  | 12 | BJ120847 BJ120847 |
| 10 | 44.2 | 1.9 | 540  | 9  | AU201678 AU201678 |
| 11 | 44.2 | 1.9 | 560  | 14 | CB401025 CB401025 |
| 12 | 44.2 | 1.9 | 584  | 12 | BJ121613 BJ121613 |
| 13 | 44.2 | 1.9 | 586  | 12 | BJ121635 BJ121635 |
| 14 | 44.2 | 1.9 | 597  | 9  | AU207635 AU207635 |
| 15 | 44.2 | 1.9 | 603  | 12 | BJ101145 BJ101145 |
| 16 | 44.2 | 1.9 | 608  | 9  | AU204573 AU204573 |
| 17 | 44.2 | 1.9 | 619  | 12 | BJ117235 BJ117235 |
| 18 | 44.2 | 1.9 | 634  | 12 | BJ124022 BJ124022 |
| 19 | 44.2 | 1.9 | 655  | 12 | BJ103250 BJ103250 |
| 20 | 43.8 | 1.9 | 1201 | 13 | BX381961 BX381961 |
| 21 | 43.4 | 1.9 | 482  | 12 | BJ101110 BJ101110 |
| 22 | 43   | 1.9 | 38   | 9  | AU109666 AU109666 |
| 23 | 42.6 | 1.9 | 631  | 12 | BJ117453 BJ117453 |
| 24 | 42.4 | 1.9 | 850  | 13 | BU348455 BU348455 |
| 25 | 42   | 1.9 | 362  | 14 | CA926583 CA926583 |
| 26 | 42   | 1.9 | 1201 | 13 | BX376097 BX376097 |
| 27 | 40.8 | 1.8 | 999  | 29 | CNS060LH CNS060LH |
| 28 | 40.6 | 1.8 | 310  | 13 | BU821300 BU821300 |
| 29 | 40.6 | 1.8 | 411  | 9  | A1102326 A1102326 |
| 30 | 40.6 | 1.8 | 479  | 14 | CA934750 CA934750 |
| 31 | 40.6 | 1.8 | 1376 | 29 | AG176243 AG176243 |
| 32 | 39.4 | 1.7 | 284  | 9  | A1101650 A1101650 |
| 33 | 39.4 | 1.7 | 456  | 9  | AA943462 AA943462 |
| 34 | 39.4 | 1.7 | 462  | 12 | BI300778 BI300778 |
| 35 | 39.4 | 1.7 | 511  | 10 | BG378762 BG378762 |
| 36 | 39.4 | 1.7 | 560  | 12 | BI296551 BI296551 |
| 37 | 39.2 | 1.7 | 404  | 12 | BJ305833 BJ305833 |
| 38 | 39.2 | 1.7 | 410  | 12 | BI611075 BI611075 |
| 39 | 39.2 | 1.7 | 437  | 13 | BY467572 BY467572 |
| 40 | 39.2 | 1.7 | 501  | 14 | CA614930 CA614930 |
| 41 | 39.2 | 1.7 | 502  | 12 | BI368223 BI368223 |
| 42 | 39.2 | 1.7 | 554  | 9  | A1457056 A1457056 |
| 43 | 39.2 | 1.7 | 564  | 12 | BI566474 BI566474 |
| 44 | 39.2 | 1.7 | 622  | 12 | BI627861 BI627861 |
| 45 | 39.2 | 1.7 | 623  | 12 | BI588865 BI588865 |

#### ALIGNMENTS

RESULT 1  
LOCUS BX381961 1201 bp mRNA linear EST 08-MAY-2003  
DEFINITION BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1072YF05 3-PRIME, mRNA sequence.

ACCESSION BX381961 GI:30453007  
VERSION BX381961  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen Contact: Feng Liang Email: fliang@lifetech.com URL:  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID: CS0D1072C03NP1.

FEATURES  
source  
1..1201  
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS010721F05"
/issue="C50D10721F05"
/issue_id="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-clio(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      95 a      191 c      115 g      55 t      745 others
ORIGIN

```

```

Query Match      2.3%; Score 51.6; DB 13; Length 1201;
Best Local Similarity 5.9%; Pred. No. 0.0025;
Matches 49; Conservative 288; Mismatches 488; Indels 2; Gaps 1;

Qy 1217 GGATAGAGAGATCTTCTGCGGACACATCTCTCGGTGTAACATGAGATTTCCTCA 1276
    |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 351 GGNANKKAAANNKKKACNNCCNMAKCCNNACKANNKKKKKACNANNNNKKCMM 410

Qy 1277 TGATTGGACCTTCAAGCTAATGGGAATAGATCTGGCGGGGTTTGTCAGTTTGG 1336
    :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 411 NNNKKKKKCNKKKKKNNKKKKKKKKKKKKNNNNNNNNNNNNNNNNNNNNNNNTK 470

Qy 1337 AAAGCGGTTATTGATCTCCGCTTGTCATCAAGATATGAG--AAATCAGCG 1394
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 471 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 530

Qy 1395 GATGTGCGCTGAAGAATCTCAGAACTTCACGTCGATCGCATCTGAAGTGTAAAG 1454
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 531 CNKMMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 590

Qy 1455 TGTGTCTGAGCCAGCGCATATGCCATGTTCAAAGTCAAGGCGGATTCAGAGAAAAG 1514
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 591 KNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 650

Qy 1515 AAAAATAAAGTAAAGCTTAAGCGGGATATCGAATTATGATAAGGTGGTGCAC 1574
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 651 MNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 710

Qy 1575 ATCTGAGCTGCAATATCAACTCAGGCACTTGCTGATGATGCAATATTTTCA 1634
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 711 KTKKKKKKNNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 770

Qy 1635 GGCACCACTGAACCAAGCGGTGATAAAGCATATGACAGATGCTCAAACTTTCT 1694
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 771 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 830

Qy 1695 GATGACGAAAGAAATTTGGTTAGACATATCTCCGCTTGTCCTCATGAGCG 1754
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 831 MKBMMKCKMKCKMKCKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMMKMM 890

Qy 1755 GATGCAAAAGCAAGTATGCTCGACTATGATCCAGAGATCGAATGGTAAAGCT 1814
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 891 MMBBKKKMMKKKMMKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 950

Qy 1815 AGTGTCTCATGATTATATGAGGAGAGCACTCCCAAGCTGTGGCGGTCCCGCAA 1874
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 951 CVMKMKMKKCCCKVMKMKCAKKKKMKCMCVKMKCMCKMKCMCKMKMKVKCCMKMM 1010

Qy 1875 AAAAGACGGAATTATGTCTGCTGGCGGAGCGCAATTTCAGATCTTCCCGGCTTGGCCA 1934
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1011 MNNNNCAKMMMBMMMBMMBAVVMMNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1070

Qy 1935 GCACCTATTTCTGCTGCGCTGATAGACCAAAATGTTATCAATGATTTGGCTTAC 1994
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1071 VCKRMBKCMKSGCMGCMCKGCGVGCMSCCMMBSKGCCKKMBKCMCMCKMCECMC 1130

Qy 1995 AGACGAAATGCGCGGAGGCTTCAAATCAACCGCGTGTGAGG 2041
    :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1131 BKVCVKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1177

```

```

RESULT 2
AU239500
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITILE
JOURNAL
COMMENT
Unpublished
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluscript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source
1..618
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF19-89-D06"
/tissue_type="mixture of silique and flower"
/lab_host="DH10B"
/clone_id="RAF19"
/note="Site 1: BamHI; Site 2: SalI; Subtraction Library"

BASE COUNT      158 a      143 c      148 g      166 t      3 others
ORIGIN

```

```

Query Match      2.2%; Score 49.8; DB 9; Length 618;
Best Local Similarity 60.9%; Pred. No. 0.0072;
Matches 81; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 722 TCATTGGCGCTGCGATTTCCGACTCGTGTGCAACGAAGCTTCATGCTGGGCTAG 781
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 82 TCATTGGCGCTGGAATCTCCGGAATATCGGCGCGAAGATTCTGTCGAGAACGGAGTAG 141

Qy 782 ACGATGTACAATATATGAAGCAAGTATGTTTGAAGCAAGCTTGTGTCATGCTT 841
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 142 AAGATGATTGATATCTGGAAGCAAGATCGATCGAGAGAAATATCATTAACAGAAATT 201

Qy 842 TCAGGACGCTCC 854
    |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 202 TCGGTGACGTGCC 214

```

```

RESULT 3
BU123611
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
Caenorhabditis elegans

```

ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Pelodierinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 564)  
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished  
COMMENT Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers

FEATURES  
source  
1. 564  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1302a06"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
/clone\_1ib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"

BASE COUNT 170 a 123 c 115 g 151 t 5 others  
ORIGIN

Query Match 2.0%; Score 44.8; DB 12; Length 564;  
Best Local Similarity 56.6%; Pred. No. 0.2;  
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAGTGGCGGTGATGCGCTGCGATTTCCGACTCGTGTGCAAAAGCAACTG 765  
DB 244 AAACCGTCTATGACATCGTCGCGCGCGGTATTTCCGACTATCGACCGCGCTGACTG 303  
QY 766 CTTCATGCTGGGGTACAGCATGTATCAATATGAAAGCAAGTATCTGTTGAGGCAAG 825  
DB 304 ATTGAACTTGGAATGACGATTTGATATCTACGAAGTCTCGACCGATCGAGGAGCANA 363  
QY 826 CTTTGTGACATGCTTTACGAGCG 850  
DB 364 ATTCATGCCATACCGTACAGGATG 388

RESULT 4  
BU101485 481 bp mRNA linear EST 18-JAN-2002  
LOCUS BU101485 unpublished oligo-capped cDNA library, C. elegans L1 stage  
DEFINITION Caenorhabditis elegans cDNA clone yk1027a04 5', mRNA sequence.  
ACCESSION BU101485  
VERSION BU101485.1 GI:18244155  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Pelodierinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 481)  
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished  
COMMENT Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers

FEATURES  
source  
1. 481

/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1027a04"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
/clone\_1ib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"

BASE COUNT 135 a 107 c 104 g 135 t  
ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 481;  
Best Local Similarity 56.6%; Pred. No. 0.29;  
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAGTGGCGGTGATGCGCTGCGATTTCCGACTCGTGTGCAAAAGCAACTG 765  
DB 264 AAACCGTCTATGACATCGTCGCGCGCGGTATTTCCGACTATCGACCGCGCTGACTG 323  
QY 766 CTTCATGCTGGGGTACAGCATGTATCAATATGAAAGCAAGTATCTGTTGAGGCAAG 825  
DB 324 ATTGAACTTGGAATGACGATTTGATATCTACGAAGTCTCGACCGATCGAGGAGCAGA 383  
QY 826 CTTTGTGACATGCTTTACGAGCG 850  
DB 384 ATTCATGCCATACCGTACAGGATG 408

RESULT 5  
BU103567 486 bp mRNA linear EST 18-JAN-2002  
LOCUS BU103567 unpublished oligo-capped cDNA library, C. elegans L1 stage  
DEFINITION Caenorhabditis elegans cDNA clone yk1051f01 5', mRNA sequence.  
ACCESSION BU103567  
VERSION BU103567.1 GI:18246237  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea  
; Rhabditidae; Pelodierinae; Caenorhabditis.  
REFERENCE 1 (bases 1 to 486)  
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
TITLE A complementary view of the C.elegans genome  
JOURNAL Unpublished  
COMMENT Contact: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers

FEATURES  
source  
1. 486  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1051f01"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
/clone\_1ib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"

BASE COUNT 137 a 104 c 114 g 130 t 1 others  
ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 486;  
Best Local Similarity 56.6%; Pred. No. 0.29;  
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAAGTGGCGGTCATTTGGCGCTGCGATTTCCGACTCGTGGTGGCAAGCAACTG 765  
|||  
Db 193 AAACCGTCTATACATCGTCGCGCGCGGATTTCCGACTATACGACCCGCGCTCACTG 252  
|||  
QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGGAGGCAAG 825  
|||  
Db 253 ATTGAACCTTGGAATTTGACGATTTTGATATCTACGAAGGTCTCGACCGGATCGAGGCGAGA 312  
|||  
QY 826 CTTTGGTCACATGCTTTCAGGAGC 850  
|||  
Db 313 ATTCAATGCCATACCGTACAGGATG 337  
|||

RESULT 6  
AU199996 508 bp mRNA linear EST 17-JUL-2001  
LOCUS AU199996 unpublished oligo-capped cDNA library, stage L2  
DEFINITION Caenorhabditis elegans cDNA clone yk774a07 5', mRNA sequence.  
ACCESSION AU199996  
VERSION AU199996.1 GI:14827576  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 508)  
Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
A complementary view of the C.elegans genome  
Unpublished  
Contact: Yuji Kohara  
Genome Biology Lab.  
National Institute of Genetics  
Yata 1111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykoha@lab.nig.ac.jp.  
Location/Qualifiers  
FEATURES  
source  
1..508  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk774a07"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L2"  
/clone\_lib="unpublished oligo-capped cDNA library, stage  
L2"

BASE COUNT 148 a 109 c 116 g 135 t  
ORIGIN

Query Match 1.9%; Score 44.2; DB 9; Length 508;  
Best Local Similarity 56.6%; Pred. No. 0.3;  
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAAGTGGCGGTCATTTGGCGCTGCGATTTCCGACTCGTGGTGGCAAGCAACTG 765  
|||  
Db 205 AAACCGTCTATACATCGTCGCGCGCGGATTTCCGACTATACGACCCGCGCTCACTG 264  
|||  
QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGGAGGCAAG 825  
|||  
Db 265 ATTGAACCTTGGAATTTGACGATTTTGATATCTACGAAGGTCTCGACCGGATCGAGGCGAGA 324  
|||  
QY 826 CTTTGGTCACATGCTTTCAGGAGC 850  
|||  
Db 325 ATTCAATGCCATACCGTACAGGATG 349  
|||

RESULT 7  
BU108542 511 bp mRNA linear EST 30-MAY-2003  
LOCUS BU108542 unpublished oligo-capped cDNA library, C. elegans L1 stage  
DEFINITION

ACCESSION BU108542  
VERSION BU108542.2 GI:31243973  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 511)  
Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
A complementary view of the C.elegans genome  
Unpublished  
Contact: Tadao Shin-i  
On Jan 23, 2002 this sequence version replaced gi:18268570.  
Center for Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
FEATURES  
source  
1..511  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1109b04"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
/clone\_lib="unpublished oligo-capped cDNA library, C.  
elegans L1 stage"

BASE COUNT 150 a 108 c 117 g 135 t 1 others  
ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 511;  
Best Local Similarity 56.6%; Pred. No. 0.3;  
Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAAGTGGCGGTCATTTGGCGCTGCGATTTCCGACTCGTGGTGGCAAGCAACTG 765  
|||  
Db 209 AAACCGTCTATACATCGTCGCGCGCGGATTTCCGACTATACGACCCGCGCTCACTG 268  
|||  
QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGGAGGCAAG 825  
|||  
Db 269 ATTGAACCTTGGAATTTGACGATTTTGATATCTACGAAGGTCTCGACCGGATCGAGGCGAGA 328  
|||  
QY 826 CTTTGGTCACATGCTTTCAGGAGC 850  
|||  
Db 329 ATTCAATGCCATACCGTACAGGATG 353  
|||

RESULT 8  
AU203802 513 bp mRNA linear EST 30-MAY-2003  
LOCUS AU203802 unpublished oligo-capped cDNA library, stage L1  
DEFINITION Caenorhabditis elegans cDNA clone yk825a11 5', mRNA sequence.  
ACCESSION AU203802  
VERSION AU203802.2 GI:31231885  
KEYWORDS EST.  
SOURCE Caenorhabditis elegans  
ORGANISM Caenorhabditis elegans  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea  
; Rhabditidae; Peloderinae; Caenorhabditis.  
1 (bases 1 to 513)  
Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.  
and Sugano,S.  
A complementary view of the C.elegans genome  
Unpublished  
On Jul 17, 2001 this sequence version replaced gi:14834726.  
Contact: Yuji Kohara  
Genome Biology Lab.



Db 349 ATTGATGCATACCGTACAGATG 373

RESULT 11

LOCUS CB401025 560 bp mRNA linear EST 15-MAY-2003

DEFINITION OSTF186f11\_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.

ACCESSION CB401025

VERSION CB401025.1 GI:30742752

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS 1. Rhabditidae; Peloderinae; Caenorhabditis.

2. (bases 1 to 560)

Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacquot, L., Bertin, N., Janky, R., Moore, T., Hudson, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S., Endress, G.A., Jena, S., Chevet, E., Papasotiropoulos, V., Tollas, P.P., Placek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H., Doucette-Stamm, L., Hill, D.E. and Vidal, M.

C. elegans ORFome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression

Nat. Genet., (2003) In press

CONTACT: Vidal, M

Journal Comment: Marc Vidal Laboratory Dana Farber Cancer Institute 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA

Tel: 617 632 5180

Fax: 617 632 5739

Email: Marc.Vidal@dfci.harvard.edu

Designated tag of Gateway entry clones. The primers used were designed on the predicted protein encoding ORF. C. elegans ORFome cloning project : Contact david\_hill@dfci.harvard.edu or marc\_vidal@dfci.harvard.edu

POLYA=No.

FEATURES

source Location/Qualifiers

1..560

/organism="Caenorhabditis elegans"

/mol\_type="mRNA"

/strain="N2"

/db\_xref="taxon:6239"

/sex="Hermaphrodite and male"

/tissue\_type="whole animal"

/dev\_stage="mixed stage"

/clone\_1lb="AD-wrmcDNA"

/note="The AD-wrmcDNA library was generated with poly(A)+ RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauers and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pRC6"

BASE COUNT 175 a 116 c 122 g 147 t

ORIGIN

Query Match 1.9%; Score 44.2; DB 14; Length 560;

Best Local Similarity 56.6%; Pred. No. 0.3;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AAGCCGAAGTGGCGGTCATTTGGCGGTCGATTTCCGAGCTCGTGGCAAGCAACTG 765

DB 69 AAACCGTATAGCAATCGTCGGCGCGGATTTCCGAGCTATGACCGCGCGTCGACTG 128

QY 766 CTTCATGCTGGGAGACGATGTACATATATGAAACAGATGATCTGTGGAGGCAAG 825

DB 129 ATTGAACCTTGAAATGACGATTTGATATCTACGAAGGCTCTGACCGGATCGAGGAGA 168

QY 826 CTTCGTCACATGCTTCAGGACG 850

DB 189 ATTATGCAATACCGTACAGATG 213

RESULT 12

Bu121613

LOCUS Bu121613 584 bp mRNA linear EST 23-JAN-2002

DEFINITION Bu121613 unpublished oligo-capped cDNA library, C. elegans L1 stage

ACCESSION Caenorhabditis elegans cDNA clone yk1278f01 5', mRNA sequence.

VERSION Bu121613

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS 1. Rhabditidae; Peloderinae; Caenorhabditis.

2. (bases 1 to 584)

Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.

A complementary view of the C. elegans genome

Journal Comment: Unpublished

CONTACT: Tadao Shin-i

Journal Comment: Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp.

FEATURES

source Location/Qualifiers

1..584

/organism="Caenorhabditis elegans"

/mol\_type="mRNA"

/strain="N2"

/db\_xref="taxon:6239"

/clone\_1lb="yk1278f01"

/sex="hermaphrodite"

/tissue\_type="whole animal"

/dev\_stage="L1"

/clone\_1lb="unpublished oligo-capped cDNA library, C. elegans L1 stage"

BASE COUNT 178 a 122 c 130 g 154 t

ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 584;

Best Local Similarity 56.6%; Pred. No. 0.31;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AAGCCGAAGTGGCGGTCATTTGGCGGTCGATTTCCGAGCTCGTGGCAAGCAACTG 765

DB 210 AAACCGTATAGCAATCGTCGGCGCGGATTTCCGAGCTATGACCGCGCGTCGACTG 269

QY 766 CTTCATGCTGGGAGACGATGTACATATATGAAACAGATGATCTGTGGAGGCAAG 825

DB 270 ATTGAACCTTGAAATGACGATTTGATATCTACGAAGGCTCTGACCGGATCGAGGAGA 329

QY 826 CTTCGTCACATGCTTCAGGACG 850

DB 330 ATTATGCAATACCGTACAGATG 354

RESULT 13

LOCUS Bu121613 586 bp mRNA linear EST 23-JAN-2002

DEFINITION Bu121613 unpublished oligo-capped cDNA library, C. elegans L1 stage

ACCESSION Caenorhabditis elegans cDNA clone yk1278h02 5', mRNA sequence.

VERSION Bu121613

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

AUTHORS 1. Rhabditidae; Peloderinae; Caenorhabditis.

2. (bases 1 to 586)

Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.

A complementary view of the C. elegans genome

Journal Comment: Unpublished

CONTACT: Tadao Shin-i

Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers

## FEATURES

## source

1. 586  
/organism="Caenorhabditis elegans"  
/mol\_type="mRNA"  
/strain="N2"  
/db\_xref="taxon:6239"  
/clone="yk1278h02"  
/sex="hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="L1"  
/clone\_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"  
BASE COUNT 179 a 123 c 130 g 154 t  
ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 586;  
Best Local Similarity 56.6%; Pred. No. 0.31;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAGTGGCGGTCATTGGCGCTGCGACTTCGAGTGGCAAGCAACTG 765

DB 210 AAACCGCTATAGCAATCGTCGGCGCGGTATTTCCGACTATCGACCGCGCTGACTG 269

QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGGAGGCAAG 825

DB 270 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGGAGCA 329

QY 826 CTTTGTGCATGCTTTTCAGGAGC 850

DB 330 ATTCAATCCATACCGTACCAAGATG 354

RESULT 14  
AUC07635 597 bp mRNA linear EST 17-JUL-2001

LOCUS AUC07635 unpublished oligo-capped cDNA library, stage L1

DEFINITION Caenorhabditis elegans cDNA clone yk1004b11 5', mRNA sequence.

ACCESSION AUC07635

VERSION AUC07635.1 GI:14841711

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea

AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.

1 (bases 1 to 597)

Rhabditidae; Peioderinae; Caenorhabditis.

1 (bases 1 to 597)

A complementary view of the C.elegans genome

Unpublished

CONTACT: Yuji Kohara

Genome Biology Lab.

National Institute of Genetics

Yata 1111, Mishima, Shizuoka 411, Japan

Tel: 81-559-81-6854

Fax: 81-559-81-6855

Email: ykohara@lab.nig.ac.jp.

Location/Qualifiers

1. 597

/organism="Caenorhabditis elegans"

/mol\_type="mRNA"

/strain="N2"

/db\_xref="taxon:6239"

/clone="yk1004b11"

/sex="Hermaphrodite"

/tissue\_type="whole animal"

/dev\_stage="L1"

/clone\_lib="unpublished oligo-capped cDNA library, stage

BASE COUNT 179 a 126 c 131 g 161 t  
ORIGIN

Query Match 1.9%; Score 44.2; DB 9; Length 597;

Best Local Similarity 56.6%; Pred. No. 0.31;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAGTGGCGGTCATTGGCGCTGCGACTTCGAGTGGCAAGCAACTG 765

DB 223 AAACCGCTATAGCAATCGTCGGCGCGGTATTTCCGACTATCGACCGCGCTGACTG 282

QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGGAGGCAAG 825

DB 283 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGGAGCA 342

QY 826 CTTTGTGCATGCTTTTCAGGAGC 850

DB 343 ATTCAATCCATACCGTACCAAGATG 367

RESULT 15  
BJ101145 603 bp mRNA linear EST 18-JAN-2002

LOCUS BJ101145 unpublished oligo-capped cDNA library, C. elegans L1 stage

DEFINITION Caenorhabditis elegans cDNA clone yk1022g05 5', mRNA sequence.

ACCESSION BJ101145

VERSION BJ101145.1 GI:18243815

KEYWORDS EST.

SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea

AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.

1 (bases 1 to 603)

Rhabditidae; Peioderinae; Caenorhabditis.

1 (bases 1 to 603)

A complementary view of the C.elegans genome

Unpublished

CONTACT: Tadao Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 603

/organism="Caenorhabditis elegans"

/mol\_type="mRNA"

/strain="N2"

/db\_xref="taxon:6239"

/clone="yk1022g05"

/sex="hermaphrodite"

/tissue\_type="whole animal"

/dev\_stage="L1"

/clone\_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

BASE COUNT 180 a 130 c 131 g 162 t

ORIGIN

Query Match 1.9%; Score 44.2; DB 12; Length 603;

Best Local Similarity 56.6%; Pred. No. 0.31;

Matches 82; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 706 AACCCGAAGTGGCGGTCATTGGCGCTGCGACTTCGAGTGGCAAGCAACTG 765

DB 229 AAACCGCTATAGCAATCGTCGGCGCGGTATTTCCGACTATCGACCGCGCTGACTG 288

QY 766 CTTTCATGCTGGGGTAGACGATGTTACATATATGAAGCAAGTATCGTGTGGAGGCAAG 825

DB 289 ATTGAACCTTGGAATTGACGATTTTGATATCTACGAAGTCTCGACCGGATCGAGGAGCA 348

QY 826 CTTTGTGCATGCTTTTCAGGAGC 850

Mon Nov 24 16:22:33 2003

us-09-434-837-10.rst

Page 8

Db 349 ATTGATGCCATACCGTACAGAGATG 373

Search completed: November 24, 2003, 04:01:39  
Job time : 4835.41 secs